

From: Chan, Christina
Sent: Tuesday, January 14, 2003 3:06 PM
To: Holleran, Anne; STIC-Biotech/ChemLib
Subject: RE: RUSH sequence search for 09/506,079
Importance: High

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

Chris

-----Original Message-----

From: Holleran, Anne
Sent: Tuesday, January 14, 2003 12:12 PM
To: Chan, Christina
Subject: RUSH sequence search for 09/506,079

Please approve and forward to STIC the following RUSH sequence search request. This is for an amendment due this biweek. Thanks.

Please search the following for 09/506,079:

commercial and interference databases oligomer search of SEQ ID NO: 1(aa)

commercial and interference databases oligomer search of SEQ ID NO: 2(aa)

SEQ ID NO: 1 is a fragment of SEQ ID NO: 2.

Anne Holleran
AU: 1642
Tel: 308-8892
RM: 8e03
mailbox: 8e12

Point of Contact:
Barb O'Bryen
Technical Information Specialist
STIC CM1 6A05 308-4291

Searcher: BOB
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 1-15-03 ✓
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:11:49 ; Search time 5.0763 Seconds
(without alignments)
457.894 Million cell updates/sec

Title: US-09-506-079E-1

Perfect score: 1 GKHXXXPRPAVPYKRXKP.....VGRGXDPDAHVAVXLSRYEG 79

Sequence: OLIGO

Scoring table: Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCrus.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	21.5	79	4	US-09-630-155-1
2	17	21.5	419	4	US-09-630-155-2
3	7	8.9	635	2	US-09-014-969-11
4	6	7.6	12	2	US-08-811-492-146
5	6	7.6	15	1	US-08-036-555B-49
6	6	7.6	15	1	US-08-469-569-49
7	6	7.6	15	1	US-08-249-322A-49
8	6	7.6	15	1	US-08-469-569-49
9	6	7.6	15	2	US-08-734-591A-49
10	6	7.6	15	2	US-08-469-660-49
11	6	7.6	15	3	US-08-341-018-82
12	6	7.6	15	4	US-08-470-335-49
13	6	7.6	15	4	US-08-735-021-49
14	6	7.6	15	4	US-08-734-664A-49
15	6	7.6	15	4	US-08-470-339-49
16	6	7.6	15	4	US-08-467-602-49
17	6	7.6	15	5	PCT-US94-05083C-46
18	6	7.6	15	5	PCT-US94-05083C-46
19	6	7.6	15	5	PCT-US94-05083C-46
20	6	7.6	15	1	US-08-469-569-37
21	6	7.6	15	1	US-08-249-322A-37
22	6	7.6	15	1	US-08-469-526A-37
23	6	7.6	15	2	US-08-734-591A-37
24	6	7.6	15	2	US-08-469-660-37
25	6	7.6	15	4	US-08-470-335-37
26	6	7.6	15	4	US-08-735-021-37
27	6	7.6	15	4	US-08-734-664A-37

28	6	7.6	16	4	US-08-470-339-37	Sequence 37, Appl
29	6	7.6	16	4	US-08-467-602-37	Sequence 37, Appl
30	6	7.6	16	5	PCT-US94-05083C-37	Sequence 37, Appl
31	6	7.6	16	5	PCT-US94-05083C-37	Sequence 37, Appl
32	6	7.6	34	1	US-08-118-270-81	Sequence 81, Appl
33	6	7.6	34	5	PCT-US93-08528-81	Sequence 81, Appl
34	6	7.6	74	4	US-09-134-001C-4092	Sequence 15, Appl
35	6	7.6	145	1	US-08-565-36-15	Sequence 15, Appl
36	6	7.6	166	4	US-09-615-192A-273	Sequence 27, Appl
37	6	7.6	290	4	US-09-386-653A-7	Sequence 7, Appl
38	6	7.6	304	4	US-09-088-651-2	Sequence 2, Appl
39	6	7.6	349	1	US-08-872-784-4	Sequence 4, Appl
40	6	7.6	349	2	US-09-100-851-4	Sequence 4, Appl
41	6	7.6	374	1	US-08-265-294-4	Sequence 14, Appl
42	6	7.6	374	1	US-08-095-726-14	Sequence 14, Appl
43	6	7.6	374	1	US-08-096-623A-14	Sequence 14, Appl
44	6	7.6	382	1	US-07-783-705A-3	Sequence 3, Appl
45	6	7.6	422	4	US-09-625-188-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-630-155-1
; Sequence 1, Application US/09630155
; Patent No. 6414130

GENERAL INFORMATION:

APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:

ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square

CITY: Seattle
STATE: Washington

COUNTRY: U.S.A.
ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible

OPERATING SYSTEM: Windows95
SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309

REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:

TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 79
STRANDEDNESS: Single

TOPOLOGY: unknown
MOLECULE TYPE: HER-2 ECD antagonist

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-630-155-1

Query Match 21.5%, Score 17; DB 4; Length 79;
Best Local Similarity 100.0%; Pred. No. 5e-10;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 37 VSAFSLPLAPLSPSV 53
Db 37 VSAFSLPLAPLSPSV 53

RESULT 2
US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 6414130
GENERAL INFORMATION:
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2
Query Match 21.5%; Score 17; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 VSAFYSPLAPLSPTSV 53
DB 377 VSAFYSPLAPLSPTSV 393
RESULT 3
US-09-014-969-11
Sequence 11, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-014-969-11
Query Match 8.9%; Score 7; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 APLSPTS 52
DB 9 APLSPTS 15
RESULT 4
US-08-811-492-146
Sequence 146, Application US/08811492
Patent No. 5834247
GENERAL INFORMATION:
APPLICANT: COMB, DONALD G.
APPLICANT: PERLER, FRANCINE B.
APPLICANT: JACK, WILLIAM E.
APPLICANT: XU, MING-QUN
APPLICANT: HODGES, ROBERT A.
APPLICANT: NOREN, CHRISTOPHER J.
APPLICANT: CHONG, SHAO-RONG S.C.
APPLICANT: ADAM, ERIC
TITLE OF INVENTION: MODIFIED PROTEINS, METHODS OF THEIR
TITLE OF INVENTION: PRODUCTION AND METHODS FOR PURIFICATION OF TARGET
NUMBER OF SEQUENCES: 155
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.
STREET: 32 TOZER ROAD
CITY: BEVERLY
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,492
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/580,555
FILING DATE: 29-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/496,247

FILING DATE: 28-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/146,885
FILING DATE: 03-NOV-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/004,139
FILING DATE: 09-DEC-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Gregory D
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-036C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-927-5054
TELEFAX: 509-927-1705
TELEX:
INFORMATION FOR SEQ ID NO: 146:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-811-492-146

Query Match 7.6%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 LAPISP 50
|||||
Db 7 LAPISP 12

RESULT 5
US-08-036-555B-49
Sequence 49, Application US/08036555B
Patent No. 5530109
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
APPLICANT: Chen, Maio Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
TITLE OF INVENTION: Preparation and Use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,555B
FILING DATE: 24-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 5250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-036-555B-49

Query Match 7.6%; Score 6; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 SPVSVG 61
|||||
Db 2 SPVSVG 7

RESULT 6
US-08-469-569-49
Sequence 49, Application US/08469569
Patent No. 5606032
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
APPLICANT: Chen, Maio Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
TITLE OF INVENTION: Preparation and Use
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felle & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,569
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991

ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 5250.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-469-569-49

Query Match
Best Local Similarity 100.0%; Score 6; DB 1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 7
US-08-249-322A-49
Sequence 49, Application US/08249322A
Patent No. 5716930
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew; Stroobant, Paul;
APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
APPLICANT: Chen, Maio Su; Hiles, Ian
TITLE OF INVENTION: Glial Mitogenic Factors, Their
PREPARATION AND USE
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/249,322A
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907,138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APRIL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Tsai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: LUD 250.4

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-249-322A-49

Query Match
Best Local Similarity 100.0%; Score 6; DB 1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 8
US-08-469-526A-49
Sequence 49, Application US/08469526A
Patent No. 5792849
GENERAL INFORMATION:
APPLICANT: Goodearl, Andrew
APPLICANT: Stroobant, Paul
APPLICANT: Minghetti, Luisa
APPLICANT: Waterfield, Michael
APPLICANT: Marchionni, Mark
APPLICANT: Chen, Maio Su
APPLICANT: Hiles, Ian
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
PREPARATION AND USE
NUMBER OF SEQUENCES: 187
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,526A
FILING DATE: 06 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036,555
FILING DATE: 24-MAR-1993
APPLICATION NUMBER: 07/965,173
FILING DATE: 23-OCT-1992
APPLICATION NUMBER: 07/940,389
FILING DATE: 03-SEP-1992
APPLICATION NUMBER: 07/907,138
FILING DATE: 03-JUN-1992
APPLICATION NUMBER: 07/863,703
FILING DATE: 03-APRIL-1992
APPLICATION NUMBER: U.K. 91 07566.3
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:

LENGTH: 15
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-469-526A-49

Query Match 7.6%; Score 6; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 SPVSVG 61
|||||
Db 2 SPVSVG 7

RESULT 9
US-08-734-591A-49
Sequence 49, Application US/08734591A
Patent No. 5854220
GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew
APPLICANT: Stroobant, Paul
APPLICANT: Minghetti, Luisa
APPLICANT: Waterfield, Michael
APPLICANT: Hiles, Ian
APPLICANT: Marchionni, Mark
TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
TITLE OF INVENTION: PREPARATION AND USE
NUMBER OF SEQUENCES: 187
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible Pentium
OPERATING SYSTEM: Windows95
SOFTWARE: WordPerfect (Version 7.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/734, 591A
FILING DATE: 22-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/470, 335
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/036, 555
FILING DATE: 03-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/965, 173
FILING DATE: 23-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/940, 389
FILING DATE: 03-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/907, 138
FILING DATE: 30-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/863, 703
FILING DATE: 03-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 91 07566.3
FILING DATE: 10-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 04585/00200P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 428-0200

TELEFAX: (617) 428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
US-08-734-591A-49

Query Match 7.6%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 56 SPVSVG 61
|||||
Db 2 SPVSVG 7

RESULT 10
US-08-469-660-49
Sequence 49, Application US/08469660
Patent No. 5876973
GENERAL INFORMATION:

APPLICANT: Gwynne, David I.; Marchionni, Mark;
APPLICANT: McBurney, Robert N.
TITLE OF INVENTION: INHIBITORS OF CELL PROLIFERATION,
TITLE OF INVENTION: THEIR PREPARATION AND USE
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
ZIP: 0211-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469, 660
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/011, 396
FILING DATE: 29-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/984, 085
FILING DATE: 01-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/951, 747
FILING DATE: 25-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/927, 337
FILING DATE: 10-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04585/017004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: 200154
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 15
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-469-660-49

Query Match 7.6%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 11
US-08-341-018-82
; Sequence 82, Application US/08341018A
; Patent No. 6087323
; GENERAL INFORMATION:
; APPLICANT: Gymer, David I.
; APPLICANT: Mahanthappa, Nagesh K.
; APPLICANT: Marchionni, Mark A.
; APPLICANT: Birmingham-McDonogh, Olivia
; APPLICANT: Goldin, Stanley M.
; APPLICANT: McBurney, Robert N.
; TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF
; TITLE OF INVENTION: CELLULAR COMMUNICATION
; FILE REFERENCE: 04585/041001
; CURRENT APPLICATION NUMBER: US/08/341,018A
; CURRENT FILING DATE: 1994-11-17
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Bos taurus
US-08-341-018-82

Query Match 7.6%; Score 6; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 12
US-08-470-335-49
; Sequence 49, Application US/08470335F
; Patent No. 6147190
; GENERAL INFORMATION:
; APPLICANT: GOODEARL, ANDREW
; APPLICANT: STROOBANT, PAUL
; APPLICANT: MINGHETTI, LUISA
; APPLICANT: WATERFIELD, MICHAEL
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: CHEN, MARIO S.
; APPLICANT: HILES, IAN
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: 04585/00200B
; CURRENT APPLICATION NUMBER: US/08/470,335F
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/036,555
; EARLIER FILING DATE: 1993-03-24
; NUMBER OF SEQ ID NOS: 252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Bos taurus
US-08-470-335-49

Query Match 7.6%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||

DB 2 SPVSVG 7

RESULT 13
US-08-735-021-49
; Sequence 49, Application US/08735021B
; Patent No. 6194377
; GENERAL INFORMATION:
; APPLICANT: GOODEARL, ANDREW
; APPLICANT: STROOBANT, PAUL
; APPLICANT: MINGHETTI, LUISA
; APPLICANT: WATERFIELD, MICHAEL
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: CHEN, MARIO S.
; APPLICANT: HILES, IAN
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; FILE REFERENCE: 04585/00200L
; CURRENT APPLICATION NUMBER: US/08/735,021B
; CURRENT FILING DATE: 1996-10-22
; EARLIER APPLICATION NUMBER: 08/472,065
; EARLIER FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/036,555
; EARLIER FILING DATE: 1993-03-24
; EARLIER APPLICATION NUMBER: 07/965,173
; EARLIER FILING DATE: 1992-10-23
; EARLIER APPLICATION NUMBER: 07/940,389
; EARLIER FILING DATE: 1992-09-03
; EARLIER APPLICATION NUMBER: 07/907,138
; EARLIER FILING DATE: 1992-06-30
; EARLIER APPLICATION NUMBER: 07/863,703
; EARLIER FILING DATE: 1992-04-03
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 49
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Bos taurus
US-08-735-021-49

Query Match 7.6%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
|||||
DB 2 SPVSVG 7

RESULT 14
US-08-734-664A-49
; Sequence 49, Application US/08734664A
; Patent No. 6204241
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Mario
; APPLICANT: Hiles, Ian
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; TITLE OF INVENTION: PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB


```

; COMPUTER: IBM Compatible Pentium
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,664A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,322
; FILING DATE: 26-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 91 07566.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-734-664A-49

Query Match          7.6%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
Db 2 SPVSVG 7

```

```

; EARLIER FILING DATE: 1992-09-03
; EARLIER APPLICATION NUMBER: 07/907,138
; EARLIER FILING DATE: 1992-06-30
; EARLIER APPLICATION NUMBER: 07/863,703
; EARLIER FILING DATE: 1992-04-03
; EARLIER APPLICATION NUMBER: 91 07566.3 GB
; EARLIER FILING DATE: 1999-04-10
; NUMBER OF SEQ ID NOS: 226
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Bos taurus
; US-08-470-339-49

Query Match          7.6%; Score 6; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 SPVSVG 61
Db 2 SPVSVG 7

```

Search completed: January 14, 2003, 17:15:45
 Job time : 6.0763 secs

```

RESULT 15
US-08-470-339-49
; Sequence 49, Application US/08470339C
; Patent No. 6232286
; GENERAL INFORMATION:
; APPLICANT: GOODEARTL, ANDREW
; APPLICANT: STROOBART, PAUL
; APPLICANT: MINGHETTI, LUISA
; APPLICANT: WATERFIELD, MICHAEL
; APPLICANT: MARCHIONNI, MARK
; APPLICANT: CHEN, MARIO S.
; APPLICANT: HILES, IAN
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; FILE REFERENCE: 04585/002008
; CURRENT APPLICATION NUMBER: US/08/470,339C
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/036,555
; EARLIER FILING DATE: 1993-03-24
; EARLIER APPLICATION NUMBER: 07/940,389

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:10:43 ; Search time 46.2751 Seconds
(without alignments)
1865.663 Million cell updates/sec

Title: US-09-506-079E-2

Perfect score: 419

Sequence: 1 MELALCRWGLLALLPPGA.....VGRGXDPAHVAVXLSRYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mmc:*
9: sp_organella:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriophage:*
17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	19.8	419	4	09UK79 homo sapien
2	51	12.2	149	6	09BG66 oryctolagus
3	50	11.9	1259	6	018735 canis famli
4	9	2.1	478	11	09ESE0 ratius norv
5	9	2.1	527	13	090836 gallus gall
6	9	2.1	599	13	09PSH2 gallus gall
7	9	2.1	643	11	09SERV6 mus musculu
8	9	2.1	655	11	09WVF5 mus musculu
9	9	2.1	1209	11	09QX70 ratius norv
10	9	2.1	1210	11	09EP98 mus musculu
11	8	1.9	48	11	063365 ratius norv
12	8	1.9	63	11	09QZ16 ratius norv
13	8	1.9	63	11	09QZ15 ratius norv
14	8	1.9	277	11	008558 ratius norv
15	8	1.9	277	11	0924V3 cricetus
16	8	1.9	280	13	08UVJ2 oreochromis

17	8	1.9	341	16	091072 pseudomonas
18	8	1.9	605	11	0921P2 mus musculu
19	7	1.7	9	2	093LE4 heliobacilli
20	7	1.7	61	10	08S9S5 oryza sativ
21	7	1.7	67	12	084264 european el
22	7	1.7	114	11	09D1N5 mus musculu
23	7	1.7	117	2	08VM13 mus musculu
24	7	1.7	120	11	09JIT2 mus musculu
25	7	1.7	121	16	08YZW8 anabaena sp
26	7	1.7	124	11	09CWN9 mus musculu
27	7	1.7	124	11	09WT29 mus musculu
28	7	1.7	147	11	09D0S2 mus musculu
29	7	1.7	150	6	09BG64 oryctolagus
30	7	1.7	162	10	09FPU4 petunia inh
31	7	1.7	163	11	09JLY9 mus musculu
32	7	1.7	166	5	09NMR1 leishmania
33	7	1.7	171	16	08UBQ6 agrobacteri
34	7	1.7	174	8	09XPE3 eumeces agr
35	7	1.7	178	10	09FMC7 arabidopsis
36	7	1.7	183	16	09ACW4 streptomyces
37	7	1.7	192	17	097B43 thermoplasma
38	7	1.7	209	10	09FRR1 arabidopsis
39	7	1.7	212	4	096CM4 homo sapien
40	7	1.7	222	2	09Z3Z2 pseudomonas
41	7	1.7	222	16	098K77 rhizobium 1
42	7	1.7	223	16	0913N6 pseudomonas
43	7	1.7	241	2	09JN88 streptomyces
44	7	1.7	248	2	P95438 pseudomonas
45	7	1.7	248	16	091314 pseudomonas

ALIGNMENTS

RESULT 1	ID	Q9UK79	PRELIMINARY:	PRT:	419 AA.
AC	Q9UK79	01-MAY-2000 (TREMBLER, 13, Created)			
DT	01-MAR-2001 (TREMBLER, 16, Last sequence update)				
DT	01-JUN-2002 (TREMBLER, 21, Last annotation update)				
DE	HER-2.				
GN	HER-2.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
FX	MEDLINE=99415951; PubMed=10485918;				
RA	Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;				
RL	Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; AF177761; AAD56009.2;				
DR	InterPro; IPR000494; EGFR_L_domain.				
DR	PIfam; PF00757; Furin-like; 1.				
DR	PIfam; PF01030; Recep_L_domain; 1.				
DR	SMART; SM00261; FU; 1.				
SO	SEQUENCE 419 AA; 45472 MW; FRC1BE347E2D030C CXC64;				
Query Match	19.8%; Score 83; DB 4; Length 419;				
Best Local Similarity	100.0%; Pred. No. 5.6e-77;				
Matches	83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	199 CKGSRCKGSESDCSLRTVCAGGACGKGPLPDCCHBOCACGCTGPKHSDCLACIHF 258				
DB	199 CKGSRCKGSESDCSLRTVCAGGACGKGPLPDCCHBOCACGCTGPKHSDCLACIHF 258				

```

QY 259 NMSGICELHCPALVTYNTDTFES 281
DB 259 NMSGICELHCPALVTYNTDTFES 281

RESULT 2
Q9BG66 PRELIMINARY; PRT; 149 AA.
ID Q9BG66;
AC Q9BG66;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Receptor tyrosine kinase ErbB2 (Fragmant).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBL_TaxID-9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
RA Teten F., Fischer B.;
RT "ErbB genes and epidermal growth factor- (EGF-) like ligands in the
RT per-implantation rabbit uterus and blastocyst.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF333178; AAK14371.1;
DR InterPro: IPR002174; Furin-like.
DR SMART; SM00261; FV; 2.
KW Kinase.
FT NON_TER
FT NON_TER
SQ SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;

Query Match
Best Local Similarity 100.0%; Score 51; DB 6; Length 149;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LPTDCHEGCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 281
DB 74 LPTDCHEGCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 124

RESULT 3
Q9BG66 PRELIMINARY; PRT; 1259 AA.
ID 018735;
AC 018735;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ErbB-2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBL_TaxID-9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.;
RT "CDNA cloning of erbB-2 from canine mammary gland.";
RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008451; BAA23127.1;
DR HSP; P11362; IFCK.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; EGFR_L_domain.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr-kinase.
DR InterPro: IPR004019; YLP_molif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR ProDom; PD000001; Euk_pkinase; 1.

```

```

DR SMART; SM00261; EV; 3.
DR SMART; SM00219; TYRC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; transferase; tyrosine-protein kinase.
SQ SEQUENCE 1259 AA; 137989 MW; E37364D49C4ACD46 CRC64;

Query Match
Best Local Similarity 100.0%; Score 50; DB 6; Length 1259;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 PTDCHEGCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 281
DB 232 PTDCHEGCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 281

RESULT 4
Q9SE60 PRELIMINARY; PRT; 478 AA.
ID Q9SE60;
AC Q9SE60;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Epidermal growth factor receptor related protein.
GN ERP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBL_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-CASTRO-DUODENAL MUCOUS;
RA Yu Y., Mosher J.A., Majumdar A.P.N.;
RT "Cloning of a novel EGFR-related peptide: A putative negative
RT regulator of EGFR.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF187818; AAG17037.2;
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR002174; Furin-like.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
KW Receptor.
SQ SEQUENCE 478 AA; 53233 MW; CF873A8376C519E5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 11; Length 478;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 QCAAGCTGP 247
DB 235 QCAAGCTGP 243

RESULT 5
Q90836 PRELIMINARY; PRT; 527 AA.
ID Q90836;
AC Q90836;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE EGF/TGF-alpha receptor precursor.
GN C-ERBB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBL_TaxID-9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92123214; PubMed-1732751;
RX Flickinger T.W., Malin N.J., Kung H.-J.;

```

```

RT "An alternatively processed mRNA from the avian c-erbB gene encodes a
RT soluble, truncated form of the receptor that can block ligand-
RT dependent transfection."
RL MOL. Cell. Biol. 12:883-893(1992).
DR EMBL: M77637; AAA48759.1; -.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR002174; Furin-Like.
DR Pfam: PF00757; Furin-Like; 1.
DR SMART: SM00261; FU; 2.
KW Receptor; Signal.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 527 EGF/TGF-ALPHA RECEPTOR.
SQ SEQUENCE 527 AA; 58353 MW; 764564ABCC095298 CRC64;

Query Match 2.1%; Score 9; DB 13; Length 527;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 239 QCAAGCTGP 247
Db 245 QCAAGCTGP 253

RESULT 6
Q9PSH2 PRELIMINARY; PRT; 599 AA.
ID Q9PSH2
AC Q9PSH2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Epidermal growth factor receptor (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94020816; PubMed=8414496;
RA Callaghan T., Antczak M., Flickinger T., Raines M., Myers M.,
RA Kung H.J.;
RT "A complete description of the EGF-receptor exon structure:
RT implication in oncogenic activation and domain evolution."
RL Oncogene 8:2939-2948(1993).
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR002174; Furin-Like.
DR Pfam: PF00757; Furin-Like; 2.
DR SMART: SM00261; FU; 3.
DR SMART: SM00261; FU; 3.
SQ SEQUENCE 599 AA; 66363 MW; FEAB46D293D91BD CRC64;

Query Match 2.1%; Score 9; DB 13; Length 599;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 239 QCAAGCTGP 247
Db 393 QCAAGCTGP 401

RESULT 7
Q9ERV6 PRELIMINARY; PRT; 643 AA.
ID Q9ERV6
AC Q9ERV6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Epidermal growth factor receptor isoform 2.
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEWTAC;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
RA Sinclair C.S., Pearisall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maible N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms."
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF275366; AAG28046.1; -.
DR EMBL: AF275364; AAG28046.1; JOINED.
DR EMBL: AF275365; AAG28046.1; JOINED.
DR MGD: MGI:95294; Egr.
DR InterPro: IPR000345; Cyrc_heme_bind.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR002174; Furin-Like.
DR Pfam: PF00757; Furin-Like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR SMART: SM00261; FU; 4.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 643 AA; 71476 MW; DEF22002C84911B1 CRC64;

Query Match 2.1%; Score 9; DB 11; Length 643;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 239 QCAAGCTGP 247
Db 235 QCAAGCTGP 243

RESULT 8
Q9WVF5 PRELIMINARY; PRT; 655 AA.
ID Q9WVF5
AC Q9WVF5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Epidermal growth factor receptor (Epidermal growth factor receptor
DE isoform 3).
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RA Reiter J.L., Threadgill D.W., Danielson A.J., Schell C.,
RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
RA Maible N.J.;
RT "Alternative Transcripts from the Human and Mouse EGFR Genes Encode
RT Carboxy-Terminal Truncated Receptors."
RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEWTAC;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
RA Sinclair C.S., Pearisall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Maible N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms."
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LIVER;
RX MEDLINE=21085660; PubMed=11217851;

```

RA Kawai J., Shigaoka A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
 RA Saito T., Okazaki Y., Gojobori T., Boro H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glass C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nishikido I., Pesole G., Quackenbush J.,
 RA Schmitt L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamuya M., Lee N. H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. F.,
 RA Suzuki H., Toyooka K., Wang K. H., Weltz C., Whittaker C., Wilming L.,
 RA Wuyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsuke S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AF124513; AAD44149.1; -;
 DR EMBL: AF275366; AAG28047.1; -;
 DR EMBL: AF275364; AAG28047.1; JOINED.
 DR EMBL: AF275365; AAG28047.1; JOINED.
 DR EMBL: AK004944; BAB23688.1; -;
 DR EMBL: AK004883; BAB23641.1; -;
 DR EMBL: AK004911; BAB23662.1; -;
 DR MGD: MGI:95294; Egfr.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FU; 3.
 KW Receptor.
 SQ SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;
 Query Match 2.1%; Score 9; DB 11; Length 655;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 239 QCAAGCTGP 247
 Db 235 QCAAGCTGP 243
 RESULT 9
 O9QX70 PRELIMINARY; PRT; 1209 AA.
 AC O9QX70;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Epidermal growth factor receptor.
 GN EGFR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FISHER; TISSUE=LIVER;
 RX MEDLINE=90258888; PubMed=2342466;
 RA Petch L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,
 RA Earp H.S.;
 RT "A truncated, secreted form of the epidermal growth factor receptor is
 RT encoded by an alternatively spliced transcript in normal rat tissue.";
 RL Mol. Cell. Biol. 10:2973-2982(1990).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=FISHER; TISSUE=LIVER;
 RA Petch L.A.;
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=FISHER; TISSUE=LIVER;
 RA Guttridge K., Dawson T.L., Earp H.S.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: M37394; AAF14008.1; -;
 DR HSSP: P11362; IFGR.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00261; FU; 3.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.
 SQ SEQUENCE 1209 AA; 134891 MW; 96FEE7F6CC1B7773 CRC64;
 Query Match 2.1%; Score 9; DB 11; Length 1209;
 Best Local Similarity 100.0%; Pred. No. 4.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 239 QCAAGCTGP 247
 Db 235 QCAAGCTGP 243
 RESULT 10
 O9EP98 PRELIMINARY; PRT; 1210 AA.
 AC O9EP98;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Epidermal growth factor receptor Isoform 1.
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/101.129/SVJ, AND 129/SVETAC;
 RA Reiter J.L., Thredgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
 RA Sinclair C.S., Pearshall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Mahle N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egfr transcripts encoding truncated receptor
 RT isoforms.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Reiter J.L., Thredgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
 RA Schenl C., Pearshall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Mahle N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egfr transcripts encoding truncated receptor
 RT isoforms.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275366; AAG28045.1; -;
 DR EMBL: AF275364; AAG28045.1; JOINED.
 DR EMBL: AF275365; AAG28045.1; JOINED.
 DR EMBL: AF275367; AAG24386.1; -;
 DR HSSP: P11362; IFGR.
 DR MGD: MGI:95294; Egfr.

DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR PRINTS: PRO0109; TYRKINASE.
 DR PRODOM: PD000001; Euk_pkinase; 1.
 DR SMART: SMO0261; FU; 5.
 DR SMART: SMO0220; S_Tkc; 1.
 DR SMART: SMO0219; TYKC; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR ATP-binding; Receptor; Transferase.
 SQ SEQUENCE 1210 AA; 134840 MW; 62CD021C9DE32E18 CRC64;

Query Match 2.1%; Score 9; DB 11; Length 1210;
 Best Local Similarity 100.0%; Pred. No. 4.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 QCAAGCTGP 247
 |||||
 DB 235 QCAAGCTGP 243

RESULT 11
 063365 PRELIMINARY; PRT; 48 AA.
 AC 063365;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Transforming oncogene NEU (Fragment).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91061737; PubMed=2123292;
 RA Suen T.C., Hung M.C.;
 RT "Multiple cts- and trans-acting elements involved in regulation of the
 RT neu gene";
 RL Mol. Cell. Biol. 10:6306-6315(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91172172; PubMed=1672439;
 RA Yan D.H., Hung M.C.;
 RT "Identification and characterization of a novel enhancer for the rat
 RT neu promoter";
 RL Mol. Cell. Biol. 11:1875-1882(1991).
 DR EMBL; M61004; AAA41686.1; -
 KM Oncogene; Transforming protein.
 FT NON_TER 48
 SQ SEQUENCE 48 AA; 4966 MW; 72A0E4B1E13F1129 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 48;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 LALLPPG 19
 |||||
 DB 15 LALLPPG 22

RESULT 12
 090216 PRELIMINARY; PRT; 63 AA.

AC 090216;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE Carbonyl reductase isoform I (Fragment).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=OVARY;
 RA Espey L.L., Yoshioke S., Russell D., Ujioka T., Viadu B., Skelsey M.,
 RA Fujii S., Okamura H., Richards J.S.;
 RT "Characterization of Ovarian Carbonyl Reductase Gene Expression during
 RT Ovulation in the Gonadotropin-Primed Immature Rat";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF181955; AAF03394.1; -
 FT NON_TER 1
 SQ SEQUENCE 63 AA; 6947 MW; 03514BC1F399AA56 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 63;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LALLPPGA 20
 |||||
 DB 40 LALLPPGA 47

RESULT 13
 090215 PRELIMINARY; PRT; 63 AA.
 AC 090215;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE Carbonyl reductase isoform II (Fragment).
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=OVARY;
 RA Espey L.L., Yoshioke S., Russell D., Ujioka T., Viadu B., Skelsey M.,
 RA Fujii S., Okamura H., Richards J.S.;
 RT "Characterization of Ovarian Carbonyl Reductase Gene Expression during
 RT Ovulation in the Gonadotropin-Primed Immature Rat";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF181956; AAF03395.1; -
 FT NON_TER 1
 SQ SEQUENCE 63 AA; 6916 MW; D5964BC1E299A2F4 CRC64;

Query Match 1.9%; Score 8; DB 11; Length 63;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 LALLPPGA 20
 |||||
 DB 40 LALLPPGA 47

RESULT 14
 008558 PRELIMINARY; PRT; 277 AA.
 ID 008558;
 AC 008558;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Inducible carbonyl reductase.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTER; TISSUE=OVARY;
RX MEDLINE=97167735; PubMed=9015353;
RA Aoki H., Okada T., Mizutani T., Numata Y., Minegishi T., Miyamoto K.:
RT "Identification of two closely related genes, inducible and
RL noninducible carbonyl reductases in the rat ovary."
CC Biochem. Biophys. Res. Commun. 230:518-523(1997).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC EMBL: D89069; BAA19007.1; -.
DR HSSP: P50162; IAE1.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 277 AA; 30654 MW; DDA015D1B71673A7 CRC64;

Query Match
Best Local Similarity 1.9%; Score 8; DB 11; Length 277;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LALLPPGA 20
   |||||
Db 254 LALLPPGA 261

RESULT 15
0924V3 PRELIMINARY; PRT; 277 AA.
AC 0924V3;
DT 01-DEC-2001 (TREMUREL. 19, Created)
DT 01-DEC-2001 (TREMUREL. 19, Last sequence update)
DT 01-MAR-2002 (TREMUREL. 20, Last annotation update)
DE Carbonyl reductase 1.
GN CHCR1.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Terada T., Sugihara Y., Nakamura K.;
RT "Chinese hamster carbonyl reductase."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC EMBL: AB043541; BAB62840.1; -.
DR InterPro: IPR002198; ADH_short.
DR Pfam: PF00106; adh_short; 1.
DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 277 AA; 30498 MW; 2F7E876E2B109A31 CRC64;

Query Match
Best Local Similarity 1.9%; Score 8; DB 11; Length 277;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LALLPPGA 20
   |||||
Db 254 LALLPPGA 261

```

Search completed: January 14, 2003, 17:14:25
 Job time : 49.2751 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:14:40 ; Search time 14.3032 Seconds
(without alignments)
568.337 Million cell updates/sec

Title: US-09-506-079E-2

Perfect score: 419

Sequence: 1 MEALALCRWGLLALLPPGA.....VGRKXPDHVAVYLSRYEG 419

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 118974 seqs, 19401057 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	83	19.8	289 10 US-09-821-883-23	Sequence 23, Appl
2	83	19.8	479 10 US-09-821-883-5	Sequence 5, Appl
3	83	19.8	555 10 US-09-821-883-1	Sequence 1, Appl
4	83	19.8	564 10 US-09-821-883-3	Sequence 3, Appl
5	83	19.8	645 10 US-09-821-161-1	Sequence 1, Appl
6	83	19.8	653 9 US-09-854-356-3	Sequence 3, Appl
7	83	19.8	690 10 US-09-821-883-2	Sequence 2, Appl
8	83	19.8	697 10 US-09-821-883-4	Sequence 4, Appl
9	83	19.8	712 9 US-09-854-356-7	Sequence 7, Appl
10	83	19.8	919 9 US-09-854-356-6	Sequence 6, Appl
11	83	19.8	1255 9 US-09-769-508-2	Sequence 2, Appl
12	83	19.8	1255 9 US-09-854-356-1	Sequence 1, Appl
13	83	19.8	1255 9 US-09-930-123-2	Sequence 2, Appl
14	83	19.8	1255 10 US-09-811-123-9	Sequence 9, Appl
15	83	19.8	1255 10 US-09-811-115-3	Sequence 3, Appl
16	51	12.2	654 9 US-09-854-356-8	Sequence 8, Appl
17	51	12.2	1256 9 US-09-854-356-2	Sequence 2, Appl
18	51	12.2	1260 9 US-09-870-759-118	Sequence 118, Appl
19	41	9.8	1236 9 US-09-854-356-14	Sequence 14, Appl

20	22	5.3	22	10	US-09-466-320-19	Sequence 19, Appl
21	22	5.3	68	10	US-09-466-320-11	Sequence 11, Appl
22	15	3.6	15	10	US-09-888-721-21	Sequence 21, Appl
23	14	3.3	19	10	US-09-466-320-20	Sequence 20, Appl
24	14	3.3	65	10	US-09-466-320-12	Sequence 12, Appl
25	9	2.1	9	9	US-09-017-743C-70	Sequence 70, Appl
26	9	2.1	478	10	US-09-867-521-2	Sequence 2, Appl
27	9	2.1	1210	10	US-09-725-433-2	Sequence 2, Appl
28	8	1.9	8	9	US-09-017-743C-92	Sequence 92, Appl
29	8	1.9	9	9	US-09-017-743C-105	Sequence 105, Appl
30	7	1.7	35	9	US-09-981-876-204	Sequence 204, Appl
31	7	1.7	36	10	US-09-864-761-48790	Sequence 48790, A
32	7	1.7	61	10	US-09-864-761-44740	Sequence 44740, A
33	7	1.7	134	10	US-09-764-847-670	Sequence 670, A
34	7	1.7	146	10	US-09-893-737-222	Sequence 222, Appl
35	7	1.7	166	10	US-09-864-761-47670	Sequence 47670, A
36	7	1.7	215	9	US-10-001-876-210	Sequence 210, Appl
37	7	1.7	221	9	US-10-036-041-74	Sequence 74, Appl
38	7	1.7	221	10	US-09-799-777-68	Sequence 68, Appl
39	7	1.7	221	12	US-10-036-342-74	Sequence 74, Appl
40	7	1.7	221	12	US-10-052-586-488	Sequence 488, Appl
41	7	1.7	222	9	US-09-981-876-135	Sequence 135, Appl
42	7	1.7	237	10	US-09-925-300-1678	Sequence 1678, Ap
43	7	1.7	238	10	US-09-925-301-988	Sequence 988, App
44	7	1.7	272	10	US-09-771-161A-149	Sequence 149, App
45	7	1.7	362	10	US-09-764-864-896	Sequence 896, App

ALIGNMENTS

RESULT 1
US-09-821-883-23
Sequence 23, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Retner
APPLICANT: Vidolovic, Damir
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821, 883
PRIOR APPLICATION NUMBER: 2001-03-30
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23
LENGTH: 289
TYPE: PRT
ORGANISM: Homo sapiens
US-09-821-883-23
Query Match 19.8%; Score 83; DB 10; Length 289;
Best Local Similarity 100.0%; Pred. No. 2.7e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 199 CKSGKMGESSEDDCSLTRIVCAGGACRCKGRLPDDCCHEQCAAGCTGPKHSIDCIACIHF 258
DB 178 CKSGKMGESSEDDCSLTRIVCAGGACRCKGRLPDDCCHEQCAAGCTGPKHSIDCIACIHF 237
OY 259 NMSGICELCPALVYNTDFES 281
DB 238 NMSGICELCPALVYNTDFES 260
RESULT 2
US-09-821-883-5
Sequence 5, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Retner

APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE OF INVENTION: Cell-Based Immunotherapy
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 479
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER300*-rcm-CSF construct
US-09-821-883-5

Query Match 19.8%; Score 83; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 4.2e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 258
DB 218 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 277
QY 259 NMSGICELHCPALVYNTDTFES 281
DB 278 NMSGICELHCPALVYNTDTFES 300

RESULT 3
US-09-821-883-1
Sequence 1, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE OF INVENTION: Cell-Based Immunotherapy
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 555
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500 construct
US-09-821-883-1

Query Match 19.8%; Score 83; DB 10; Length 555;
Best Local Similarity 100.0%; Pred. No. 4.8e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 258
DB 218 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 277
QY 259 NMSGICELHCPALVYNTDTFES 281
DB 278 NMSGICELHCPALVYNTDTFES 300

RESULT 4
US-09-821-883-3
Sequence 3, Application US/09821883

Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE OF INVENTION: Cell-Based Immunotherapy
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 564
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500* construct
US-09-821-883-3

Query Match 19.8%; Score 83; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 4.8e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 258
DB 218 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 277
QY 259 NMSGICELHCPALVYNTDTFES 281
DB 278 NMSGICELHCPALVYNTDTFES 300

RESULT 5
US-09-921-161-1
Sequence 1, Application US/09921161
Patent No. US2002009062A1
GENERAL INFORMATION:
APPLICANT: Ralph, Peter
TITLE OF INVENTION: ANALYTICAL METHOD
FILE REFERENCE: GENENT.066A
CURRENT APPLICATION NUMBER: US/09/921,161
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/225,433
PRIOR FILING DATE: 2000-08-15
NUMBER OF SEQ ID NOS: 1
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 645
TYPE: PRT
ORGANISM: Homo sapiens
US-09-921-161-1

Query Match 19.8%; Score 83; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 5.4e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 258
DB 218 CKGSRGWGSSSEDCOSLTRIVCAGGCAKCGPLPTDCHEQCAAGCTGPKHSDCLACLFH 277
QY 259 NMSGICELHCPALVYNTDTFES 281
DB 278 NMSGICELHCPALVYNTDTFES 281

RESULT 6
US-09-854-356-3
Sequence 3, Application US/09854356
Patent No. US20020177567A1
GENERAL INFORMATION:

```

: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: CURRENT FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Patentl Ver. 2.1
: SEQ ID NO 3
: LENGTH: 653
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3

```

```

Query Match
Best Local Similarity 19.8%; Score 83; DB 9; Length 653;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 199 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 258
Db 199 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 258
QY 259 NHSGICELHCPALVTYNTDFES 281
Db 259 NHSGICELHCPALVTYNTDFES 281

```

```

RESULT 7
US-09-821-883-2
: Sequence 2, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Vidovic, Damir
: APPLICANT: Gradis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: CURRENT FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 690
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500-hGM-CSF construct
US-09-821-883-2

```

```

Query Match
Best Local Similarity 19.8%; Score 83; DB 10; Length 690;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 258
Db 218 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 277
QY 259 NHSGICELHCPALVTYNTDFES 281
Db 278 NHSGICELHCPALVTYNTDFES 300

```

```

RESULT 8
US-09-821-883-4
: Sequence 4, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Vidovic, Damir
: APPLICANT: Gradis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: CURRENT FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 697
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500-hGM-CSF construct
US-09-821-883-4

```

```

Query Match
Best Local Similarity 19.8%; Score 83; DB 10; Length 697;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 199 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 258
Db 218 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 277
QY 259 NHSGICELHCPALVTYNTDFES 281
Db 278 NHSGICELHCPALVTYNTDFES 300

```

```

RESULT 9
US-09-854-356-7
: Sequence 7, Application US/09854356
: Patent No. US20020177567A1
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: CURRENT FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Patentl Ver. 2.1
: SEQ ID NO 7
: LENGTH: 712
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: fusion protein
: OTHER INFORMATION: of ECD and delta PD of human HER-2/neu
US-09-854-356-7

```

```

Query Match
Best Local Similarity 19.8%; Score 83; DB 9; Length 712;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 199 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 258
Db 199 CKGSRGCGESSEDCOSLTRIVCAGGACARCKGRLPTDCCHCCAGCTGPKHSDCLACHF 258

```

QY 259 NMSGICELHCPALVTYNTDFES 281
Db 259 NMSGICELHCPALVTYNTDFES 281

RESULT 10

US-09-854-356-6
Sequence 6, Application US/09854356
Patent No. US20020177567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheyzen, Dirk
APPLICANT: Corixa Corporation
APPLICANT: Smtlklne Beecham Biologicals S. A.
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 6
LENGTH: 919
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: fusion protein
US-09-854-356-6

Query Match 19.8%; Score 83; DB 9; Length 919;
Best Local Similarity 100.0%; Pred. No. 7.5e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRGWESSEDCQSLTRTVCAAGCARGKGPPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
Db 199 CKGSRGWESSEDCQSLTRTVCAAGCARGKGPPLPTDCHEQCAAGCTGPKHSDCLACLHF 258

QY 259 NMSGICELHCPALVTYNTDFES 281
Db 259 NMSGICELHCPALVTYNTDFES 281

RESULT 11

US-09-769-508-2
Sequence 2, Application US/09769508
Patent No. US20020155527A1
GENERAL INFORMATION:
APPLICANT: STUART, SUSAN G.
APPLICANT: MONAHAN, JOHN J.
APPLICANT: LANGTON, BEATRICE CLAUDIA
APPLICANT: HANCOCK, MIRIAM E.C.
APPLICANT: CHAO, LORRINE A.
APPLICANT: BLUFORD, PETER
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
FILE REFERENCE: BEBIO-111-C1
CURRENT APPLICATION NUMBER: US/09/769,508
CURRENT FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-769-508-2

Query Match 19.8%; Score 83; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 9.8e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRGWESSEDCQSLTRTVCAAGCARGKGPPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
Db 199 CKGSRGWESSEDCQSLTRTVCAAGCARGKGPPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
QY 259 NMSGICELHCPALVTYNTDFES 281
Db 259 NMSGICELHCPALVTYNTDFES 281

RESULT 12

US-09-854-356-1
Sequence 1, Application US/09854356
Patent No. US20020177567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheyzen, Dirk
APPLICANT: Corixa Corporation
APPLICANT: Smtlklne Beecham Biologicals S. A.
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human HER-2/neu protein
NAME/KEY: DOMAIN
LOCATION: (1)..(653)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (676)..(1255)
OTHER INFORMATION: intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1255)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1048)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
US-09-854-356-1

Query Match 19.8%; Score 83; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 9.8e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRGWESSEDCQSLTRTVCAAGCARGKGPPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
Db 199 CKGSRGWESSEDCQSLTRTVCAAGCARGKGPPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
QY 259 NMSGICELHCPALVTYNTDFES 281
Db 259 NMSGICELHCPALVTYNTDFES 281

RESULT 13

US-09-930-125-2
Sequence 2, Application US/09930125
Publication No. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Kalos, Michael D.

```
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vedvick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930,125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-125-2
```

```
Query Match          19.8%; Score 83; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 9.8e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 199 CKGSRGWESESDCQSLRTVACAGCARKGKPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
Db 199 CKGSRGWESESDCQSLRTVACAGCARKGKPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
QY 259 NMSGICELHCPALVTYNTDFES 281
Db 259 NMSGICELHCPALVTYNTDFES 281
```

```
RESULT 14
US-09-811-123-9
; Sequence 9, Application US/09811123
; Patent No. US2002001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Slikowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ERBB
; FILE REFERENCE: GENEY.073A2
; CURRENT APPLICATION NUMBER: US/09/811,123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9
```

```
Query Match          19.8%; Score 83; DB 10; Length 1255;
Best Local Similarity 100.0%; Pred. No. 9.8e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 199 CKGSRGWESESDCQSLRTVACAGCARKGKPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
Db 199 CKGSRGWESESDCQSLRTVACAGCARKGKPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
QY 259 NMSGICELHCPALVTYNTDFES 281
Db 259 NMSGICELHCPALVTYNTDFES 281
```

```
RESULT 15
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. US20020035736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
```

```
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENEY.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3
```

```
Query Match          19.8%; Score 83; DB 10; Length 1255;
Best Local Similarity 100.0%; Pred. No. 9.8e-68;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 199 CKGSRGWESESDCQSLRTVACAGCARKGKPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
Db 199 CKGSRGWESESDCQSLRTVACAGCARKGKPLPTDCHEQCAAGCTGPKHSDCLACLHF 258
QY 259 NMSGICELHCPALVTYNTDFES 281
Db 259 NMSGICELHCPALVTYNTDFES 281
```

```
Search completed: January 14, 2003, 17:21:51
Job time : 16.3032 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:11:49 ; Search time 26.9237 Seconds
(without alignments)
457.894 Million cell updates/sec

Title: US-09-506-079E-2

Perfect score: 419
Sequence: 1 MEALALCRWGLLALLPPGA.....VGRCDPAHVAVXLSRYEG 419

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	19.8	419	4	US-09-630-155-2
2	83	19.8	624	3	US-08-422-108-1
3	83	19.8	624	4	US-08-422-734-1
4	83	19.8	782	2	US-09-146-283-4
5	83	19.8	782	3	US-08-579-823A-4
6	83	19.8	782	4	US-09-344-195-4
7	83	19.8	1255	1	US-08-467-083-68
8	83	19.8	1255	1	US-08-414-417B-68
9	83	19.8	1255	2	US-08-484-438-8
10	83	19.8	1255	2	US-08-486-348A-68
11	83	19.8	1255	2	US-08-625-101-2
12	83	19.8	1255	2	US-08-468-545B-68
13	83	19.8	1255	3	US-08-356-786-2
14	83	19.8	1255	3	US-08-466-680B-68
15	52	12.4	97	1	US-09-046-783-3
16	52	12.4	97	4	US-09-046-783-3
17	17	4.1	17	1	US-08-467-083-61
18	17	4.1	17	1	US-08-414-417B-61
19	17	4.1	17	2	US-08-486-348A-61
20	17	4.1	17	2	US-08-468-545B-61
21	17	4.1	17	3	US-08-466-680B-61
22	15	3.6	15	1	US-08-467-083-30
23	15	3.6	15	1	US-08-467-083-31
24	15	3.6	15	1	US-08-467-083-56
25	15	3.6	15	1	US-08-414-417B-30
26	15	3.6	15	1	US-08-414-417B-31
27	15	3.6	15	1	US-08-414-417B-56

28	15	3.6	15	2	US-08-486-348A-30	Sequence 30, Appl
29	15	3.6	15	2	US-08-486-348A-31	Sequence 31, Appl
30	15	3.6	15	2	US-08-486-348A-56	Sequence 56, Appl
31	15	3.6	15	2	US-08-468-545B-30	Sequence 30, Appl
32	15	3.6	15	2	US-08-468-545B-31	Sequence 31, Appl
33	15	3.6	15	2	US-08-468-545B-56	Sequence 56, Appl
34	15	3.6	15	3	US-08-466-680B-30	Sequence 30, Appl
35	15	3.6	15	3	US-08-466-680B-31	Sequence 31, Appl
36	15	3.6	15	3	US-08-466-680B-56	Sequence 56, Appl
37	14	3.3	79	4	US-09-630-155-1	Sequence 1, Appl
38	13	3.1	18	1	US-08-467-083-60	Sequence 60, Appl
39	13	3.1	18	1	US-08-414-417B-60	Sequence 60, Appl
40	13	3.1	18	2	US-08-486-348A-60	Sequence 60, Appl
41	13	3.1	18	2	US-08-468-545B-60	Sequence 60, Appl
42	13	3.1	18	3	US-08-466-680B-60	Sequence 60, Appl
43	10	2.4	10	3	US-08-159-339A-265	Sequence 265, App
44	10	2.4	10	3	US-08-159-339A-267	Sequence 267, App
45	10	2.4	10	3	US-08-159-339A-284	Sequence 284, App

ALIGNMENTS

RESULT 1
US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 6414130
GENERAL INFORMATION:
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2
Query Match 19.8%; Score 83; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.5e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 199 CGSKGKMSSESDOSLRTVCAAGCARGKGLPDDCCHEGCAAGCTGPKHSDDLACLHF 258
DB 199 CGSKRGWESSDDCSLRTVCAAGCARGKGLPDDCCHEGCAAGCTGPKHSDDLACLHF 258
OY 259 NMSGICELHCPALVYNTDFEES 281

DB 259 NHSGICELHCPALVYNTDTFES 281

US-08-422-108-1

Sequence 1, Application US/08422108

Patent No. 6015567

GENERAL INFORMATION:

APPLICANT: Hudziak, Robert M.

APPLICANT: Shepard, H. Michael

APPLICANT: Ullrich, Axel

TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/422,108

FILING DATE: 14-Apr-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/355460

FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/048346

FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/354319

FILING DATE: 19-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M

REGISTRATION NUMBER: 00,000

REFERENCE/DOCKET NUMBER: 554C2D2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 624 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-422-108-1

Query Match 19.8%; Score 83; DB 3; Length 624;
Best Local Similarity 100.0%; Pred. No. 2.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 178 CKGSRGWSSEDCOSLFTTVACGAGCARCKGRLPTDCHBQCAAGCTGPKHSDCLACLFH 237

QY 259 NHSGICELHCPALVYNTDTFES 281

US-08-422-734-1

Sequence 1, Application US/08422734

Patent No. 6333169

GENERAL INFORMATION:

APPLICANT: Hudziak, Robert M.

APPLICANT: Shepard, H. Michael

APPLICANT: Ullrich, Axel

TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/422,734

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/422108

FILING DATE: 14-Apr-1995

APPLICATION NUMBER: 08/355460

FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/048346

FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/354319

FILING DATE: 19-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M

REGISTRATION NUMBER: 00,000

REFERENCE/DOCKET NUMBER: 554C2D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 624 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-422-734-1

Query Match 19.8%; Score 83; DB 4; Length 624;
Best Local Similarity 100.0%; Pred. No. 2.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 178 CKGSRGWSSEDCOSLFTTVACGAGCARCKGRLPTDCHBQCAAGCTGPKHSDCLACLFH 237

QY 259 NHSGICELHCPALVYNTDTFES 281

US-09-146-283-4

Sequence 4, Application US/09146283

Patent No. 5976546

GENERAL INFORMATION:

APPLICANT: Laus, Reiner

APPLICANT: Wu, Hongyu

TITLE OF INVENTION: Immunostimulatory Compositions

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Denlinger & Associates

STREET: 350 Cambridge Ave. Suite 250

CITY: Palo Alto

RESULT 5
US-08-579-823A-4
Sequence 4, Application US/08579823A
Patent No. 6080409
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Composition and Method
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/579, 823A
FILING DATE: 03-DEC-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702

ORIGINAL SOURCE ;

ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
SEQUENCE DESCRIPTION: SEQ ID NO: 4
US-09-344-195-4

Query Match 19.8%; Score 83; DB 1; Length 782;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
259 NMSGICELHCPALVYNTDTFES 281
259 NMSGICELHCPALVYNTDTFES 281

RESULT 7
US-08-467-083-68
Sequence 68, Application US/08467083
Patent No. 5726023
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 68
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,083
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/414,417
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANBERY
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-467-083-68

Query Match 19.8%; Score 83; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
259 NMSGICELHCPALVYNTDTFES 281
259 NMSGICELHCPALVYNTDTFES 281

DB 259 NMSGICELHCPALVYNTDTFES 281

RESULT 8
US-08-414-417B-68
Sequence 68, Application US/08414417B
Patent No. 5801005
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/414,417B
FILING DATE: 31-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-414-417B-68

Query Match 19.8%; Score 83; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
259 NMSGICELHCPALVYNTDTFES 281
259 NMSGICELHCPALVYNTDTFES 281

RESULT 9
US-08-484-438-8
Sequence 8, Application US/08484438
Patent No. 5811098
Patent No. 5811098 5780031
GENERAL INFORMATION:
APPLICANT: Ploumanou, Gregory D.
APPLICANT: Culouscou, Jean-Michel
APPLICANT: Shoyab, Mohammed
APPLICANT: Stegall, Clay B.
APPLICANT: Hellstr m, Ingegerd
TITLE OF INVENTION: HELLSTR m, Karl E.
NUMBER OF SEQUENCES: 42
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 42
FILING DATE: 31-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-414-417B-68

Query Match 19.8%; Score 83; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
199 CKGSRGWSESDCQSLTRTVACAGGACRCKGPIPTDCHEQCAAGCTGPRHSDCLALHF 258
259 NMSGICELHCPALVYNTDTFES 281
259 NMSGICELHCPALVYNTDTFES 281

STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-8

Query Match 19.8%; Score 83; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGRCMWSESDCSLRTYVACAGCARCKGPLPTDCCHBQCAAGCTGPKHSDCLACLF 258
DB 199 CKGRCMWSESDCSLRTYVACAGCARCKGPLPTDCCHBQCAAGCTGPKHSDCLACLF 258

QY 259 NMSGICELHCPALVYNTDFES 281
DB 259 NMSGICELHCPALVYNTDFES 281

RESULT 10
US-08-486-348A-68
Sequence 68, Application US/08486348A
Patent No. 5846538
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match 19.8%; Score 83; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGRCMWSESDCSLRTYVACAGCARCKGPLPTDCCHBQCAAGCTGPKHSDCLACLF 258
DB 199 CKGRCMWSESDCSLRTYVACAGCARCKGPLPTDCCHBQCAAGCTGPKHSDCLACLF 258

QY 259 NMSGICELHCPALVYNTDFES 281
DB 259 NMSGICELHCPALVYNTDFES 281

RESULT 11
US-08-625-101-2
Sequence 2, Application US/08625101
Patent No. 5869445
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
NUMBER OF INVENTION: ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-625-101-2

Query Match
Best Local Similarity 100.0%; Score 83; DB 2; Length 1255;
Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRMGSSSEDCOSLFTRTVCAGGCARCKGPIPTDCHEQCAAGCTGPKHSDCLACLFH 258
|||||
DB 199 CKGSRMGSSSEDCOSLFTRTVCAGGCARCKGPIPTDCHEQCAAGCTGPKHSDCLACLFH 258
|||||

QY 259 NMSGICELHCPALVYNTDTFES 281
|||||
DB 259 NMSGICELHCPALVYNTDTFES 281
|||||

RESULT 12

US-08-468-545B-68
Sequence 68, Application US/08468545B
Patent No. 5876712

GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match
Best Local Similarity 100.0%; Score 83; DB 2; Length 1255;
Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRMGSSSEDCOSLFTRTVCAGGCARCKGPIPTDCHEQCAAGCTGPKHSDCLACLFH 258
|||||
DB 199 CKGSRMGSSSEDCOSLFTRTVCAGGCARCKGPIPTDCHEQCAAGCTGPKHSDCLACLFH 258
|||||

QY 259 NMSGICELHCPALVYNTDTFES 281
|||||
DB 259 NMSGICELHCPALVYNTDTFES 281
|||||

RESULT 13

US-08-356-786-2
Sequence 2, Application US/08356786
Patent No. 5877305

GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424

Prior Application Number:
APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-356-786-2

Query Match
Best Local Similarity 100.0%; Score 83; DB 2; Length 1255;
Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CKGSRMGSSSEDCOSLFTRTVCAGGCARCKGPIPTDCHEQCAAGCTGPKHSDCLACLFH 258
|||||
DB 199 CKGSRMGSSSEDCOSLFTRTVCAGGCARCKGPIPTDCHEQCAAGCTGPKHSDCLACLFH 258
|||||

QY 259 NMSGICELHCPALVYNTDTFES 281
|||||
DB 259 NMSGICELHCPALVYNTDTFES 281
|||||

RESULT 14

US-08-466-680B-68
Sequence 68, Application US/08466680B
Patent No. 6075122

GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle

STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,680B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-466-680B-68

Query Match 19.8%; Score 83; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 4.2e-71;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 199 CKGSRGCGESSEDCOSLTRVACAGCARKCPPLPDCGHEOCAMGCTGPKHSDCIACLHF 258
|||||
Db 199 CKGSRGCGESSEDCOSLTRVACAGCARKCPPLPDCGHEOCAMGCTGPKHSDCIACLHF 258
OY 259 NMSGICELHCPALVYNTDFES 281
|||||
Db 259 NMSGICELHCPALVYNTDFES 281

RESULT 15
US-08-421-356-3
Sequence 3, Application US/08421356
Patent No. 5783404
GENERAL INFORMATION:
APPLICANT: Koski, Raymond A.
TITLE OF INVENTION: HER-2
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Denavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/421,356
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-327
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-421-356-3

Query Match 12.4%; Score 52; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.2e-42;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 96 LQRLRVGTQLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLTE 147
|||||
Db 2 LQRLRVGTQLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLTE 53

Search completed: January 14, 2003, 17:15:48
Job time : 29.9237 secs

THIS PAGE BLANK (USPTO)

FT Misc-difference 346 /note= "Encoded by CYG"
FT /label= Unknown
FT /note= "Encoded by CYC"
FT Misc-difference 356
FT /label= Unknown
FT /note= "Encoded by CMG"
FT Misc-difference 358
FT /label= Unknown
FT /note= "Encoded by ATR"
FT Misc-difference 361
FT /label= Unknown
FT /note= "Encoded by GMC"
FT Misc-difference 376
FT /label= Unknown
FT /note= "Encoded by MTA"
FT Misc-difference 389
FT /note= "Encoded by AGC"
FT Misc-difference 394
FT /label= Unknown
FT /note= "Encoded by CST"
FT Misc-difference 404
FT /label= Unknown
FT /note= "Encoded by CYG"
FT Misc-difference 413
FT /label= Unknown
FT /note= "Encoded by SAC"
XX WO200161356-A1.
XX 23-AUG-2001.
XX 16-FEB-2001; 2001WO-US05327.
XX 16-FEB-2000; 2000US-0506079.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Clinton G, Henner WD, Evans A;
XX WPI: 2001-529334/58.
XX N-PSDB: AAD15844.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Claim 8; Page 53-54; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 10⁸. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of Intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 protein
XX containing ECDIIIA generic sequence.

Query Match	67.8%;	Score 284;	DB 22;	Length 419;
Best Local Similarity	100.0%;	Pred. No. 3.2e-259;		
Matches 24;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 126 TTPYTGASPGGLREIQDRSLTECLKGGVLIQRNPQLCYDTILMKDIFHKNNQLALFLID 185

Db 126 TTPYTGASPGGLREIQDRSLTECLKGGVLIQRNPQLCYDTILMKDIFHKNNQLALFLID 185

QY	186	TNNSRACHPESPCCKSSRCMGSESSSDCCSLIRTYCAGACACAKGPIPTDCCRCQCAAGCT	245
QY	186	TNNSRACHPCSPCKSSRCMGSESSSDCCSLIRTYCAGACACAKGPIPTDCCRCQCAAGCT	245
Db	186	TNNSRACHPCSPCKSSRCMGSESSSDCCSLIRTYCAGACACAKGPIPTDCCRCQCAAGCT	245
QY	246	GPKHSDCLACLFHNSGICELCHCPALVTYNTDFESCNPPEGRTYEGASCVTACPYNKLIS	305
Db	246	GPKHSDCLACLFHNSGICELCHCPALVTYNTDFESCNPPEGRTYEGASCVTACPYNKLIS	305
QY	306	TDVGSCTLVCPPLHNDQVTAEDTGCCCKSRPCARGKHSXXPRPAAPVYXKQPYPAHP	365
Db	306	TDVGSCTLVCPPLHNDQVTAEDTGCCCKSRPCARGKHSXXPRPAAPVYXKQPYPAHP	365
QY	366	VLSFLRPSMDVSAFYSLPLALDPTSYXISPVSYSGRGXOPDDAHVAAXLSRYRG	419
Db	366	VLSFLRPSMDVSAFYSLPLALDPTSYXISPVSYSGRGXOPDDAHVAAXLSRYRG	419
RESULT 2			
AAE09208			
XX	AAE09208	standard; Protein; 419 AA.	
AC	AAE09208;		
XX	15-NOV-2001	(first entry)	
DE	Human p68HER-2	generic protein variant 6.	
XX	HRR-2;	herstatin; antagonist; extracellular domain; ECD; Herceptin;	
KW	solid tumour; cancer; polymorphism; cytostatic; gene therapy;		
KM	p68HER-2; ECDIIIIa; variant.		
XX			
OS	Homo sapiens.		
FX	Key	Location/Qualifiers	
FH	Region	1..340	
FT		/note= "Identical to N-terminal region of p185HER-2"	
FT	Domain	341..419	
FT		/label= "ECDIIIIa_variant	
FT		/note= "Extracellular domain IIIa variant"	
FT	Misc-difference	124	
FT		/note= "Represented as Agn in the parent sequence shown	
FT		in the specification"	
FT	Misc-difference	125	
FT		/note= "Represented as Agn in the parent sequence shown	
FT		in the specification"	
FT	Misc-difference	342	
FT		/label= Unknown	
FT	Misc-difference	345	
FT		/label= Unknown	
FT	Misc-difference	346	
FT		/label= Unknown	
FT	Misc-difference	356	
FT		/label= Unknown	
FT	Misc-difference	358	
FT		/label= Unknown	
FT	Misc-difference	361	
FT		/label= Asp, Ala, Val	
FT	Misc-difference	376	
FT		/label= Unknown	
FT	Misc-difference	394	
FT		/label= Unknown	
FT	Misc-difference	404	
FT		/label= Unknown	
FT	Misc-difference	413	
FT		/label= Unknown	
XX	WO200161356-A1.		
PN	23-AUG-2001.		
PD	16-FEB-2001; 2001WO-US05327.		
PF	16-FEB-2000; 2000US-0506079.		
XX			


```

XX (UYOR-) UNIV OREGON HEALTH SCI.
PA Clinton G, Henner WD, Evans A;
PI WPI: 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11, Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 1078. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIII.
XX The ECDIIII-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIIa variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX sequence listing (AAE09181).
XX
XX Sequence 419 AA:
XX
XX Query Match 67.8%; Score 284; DB 22; Length 419;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-259;
XX Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 126 TTPVTGASPGGLRELQSLTECLKGVLIQNPOLCYQDITLMKDIFHKNNQALATLID 185
DB 126 TTPVTGASPGGLRELQSLTECLKGVLIQNPOLCYQDITLMKDIFHKNNQALATLID 185
QY 186 TNRSRACHPCSPCKGSRCKWGESSEDCOSLRTTVAGGACARCKGPLPDDCCHGQCAAGT 245
DB 186 TNRSRACHPCSPCKGSRCKWGESSEDCOSLRTTVAGGACARCKGPLPDDCCHGQCAAGT 245
QY 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDTPESCNPBGRYTFGASCVYACPYNKLIS 305
DB 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDTPESCNPBGRYTFGASCVYACPYNKLIS 305
QY 306 TDVGSCTIVCPHLNQEVTAEDGTQCEKCSKPCARGHSHXPRPAVVPYXKXOPXPAHP 365
DB 306 TDVGSCTIVCPHLNQEVTAEDGTQCEKCSKPCARGHSHXPRPAVVPYXKXOPXPAHP 365
QY 366 VLSFLRPSMDYSAFYSIPLAPLDPTSVXISPSVSGRGKXDPDAHVAVXLSRREG 419
DB 366 VLSFLRPSMDYSAFYSIPLAPLDPTSVXISPSVSGRGKXDPDAHVAVXLSRREG 419
XX
XX RESULT 3
XX AAE20348
XX ID AAE20348 standard; Protein: 419 AA.
XX
XX AAE20348;
XX
XX 18-JUN-2002 (first entry)
XX
XX Human truncated HER2 protein.
XX
XX Human; tumour; endothelial growth factor receptor; EGFR; cytosolic;
XX herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung;
XX colon; glial cell tumour; cell growth.
XX
XX Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT Misc-difference 342
FT /label- Thr, Ser
FT Misc-difference 345
FT /label- Leu, Pro
FT Misc-difference 346
FT /label- Pro, Leu
FT Misc-difference 356
FT /label- Pro, Gln
FT Misc-difference 358
FT /label- Met, Leu
FT Misc-difference 361
FT /label- Gly, Asp, Ala, Val
FT Misc-difference 376
FT /label- Leu, Ile
FT Misc-difference 394
FT /label- Pro, Arg
FT Misc-difference 404
FT /label- Pro, Leu
FT Misc-difference 413
FT /label- Asp, Asn
XX
XX WO200214470-A2.
XX
XX 21-FEB-2002.
XX
XX 14-AUG-2001; 2001WO-US25502.
XX
XX 14-AUG-2000; 2000US-0638834.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton GW;
XX
XX WPI: 2002-269185/31.
XX
XX Treating solid tumor characterized by expression of endothelial growth
XX factor receptor, involves administering recombinant herstatin that
XX binds to extracellular domain of the endothelial growth factor receptor
XX
XX
XX Claim 1: Page 78-80; 82pp; English.
XX
XX The present invention relates to a method for treating a solid tumour
XX characterised by endothelial growth factor receptor (EGFR) expression.
XX The method involves administering an agent that binds to an extracellular
XX domain (ECD) of EGFR. The invention also relates to a naturally occurring
XX inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
XX expression of herstatin with p185HER2 causes a striking reduction in cell
XX growth that corresponds with suppression of p185 autophosphorylation. The
XX method or a pharmaceutical composition is useful for treating a solid
XX tumour (selected from squamous cell carcinoma, lung carcinoma, colon
XX carcinoma and glial cell tumour) characterised by EGFR expression. The
XX present sequence is human truncated HER2 protein that lacks transmembrane
XX and intracellular domains.
XX
XX Sequence 419 AA:
XX
XX Query Match 67.8%; Score 284; DB 23; Length 419;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-259;
XX Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 126 TTPVTGASPGGLRELQSLTECLKGVLIQNPOLCYQDITLMKDIFHKNNQALATLID 185
DB 126 TTPVTGASPGGLRELQSLTECLKGVLIQNPOLCYQDITLMKDIFHKNNQALATLID 185
QY 186 TNRSRACHPCSPCKGSRCKWGESSEDCOSLRTTVAGGACARCKGPLPDDCCHGQCAAGT 245
DB 186 TNRSRACHPCSPCKGSRCKWGESSEDCOSLRTTVAGGACARCKGPLPDDCCHGQCAAGT 245
QY 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDTPESCNPBGRYTFGASCVYACPYNKLIS 305
DB 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDTPESCNPBGRYTFGASCVYACPYNKLIS 305

```

```

QY 306 TDVGSCTIVCPHLNQEVTAEADGTQRCCKSPCARGXHSXXXPPRAAVPVXRXQXPAPHP 365
PT |||||||
DB 306 TDVGSCTIVCPHLNQEVTAEADGTQRCCKSPCARGXHSXXXPPRAAVPVXRXQXPAPHP 365
QY 366 VLSFLRPSMDXVSATFSLPLAPDPTSVXISPVSVGKXDPDAHVAVXLSRYEG 419
DB 366 VLSFLRPSMDXVSATFSLPLAPDPTSVXISPVSVGKXDPDAHVAVXLSRYEG 419

RESULT 4
AAE09212
ID AAE09212 standard; Protein; 419 AA.
AC AAE09212;
XX 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 10.
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Region 1..340
FH /note="Identical to N-terminal region of p185HER-2"
FH Domain 341..419
FH /label="ECDIIIA_variant"
FH /note="Extracellular domain IIA variant"
FH Misc-difference 124
FH /note="Represented as Agn in the parent sequence shown
FH in the specification"
FH Misc-difference 125
FH /note="Represented as Agn in the parent sequence shown
FH in the specification"
FH Misc-difference 342
FH /label="Unknown"
FH Misc-difference 345
FH /label="Unknown"
FH Misc-difference 346
FH /label="Unknown"
FH Misc-difference 356
FH /label="Unknown"
FH Misc-difference 358
FH /label="Unknown"
FH Misc-difference 361
FH /label="Unknown"
FH Misc-difference 376
FH /label="Unknown"
FH Misc-difference 394
FH /label="Unknown"
FH Misc-difference 404
FH /label="Unknown"
FH Misc-difference 413
FH /label="Unknown"
FH /note="p68HER-2 generic sequence (AAE09181) Xaa
FH substituted with Asn"
XX
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
DR

```

```

XX New polypeptide, which binds to the extracellular domain of HER-2 for
PT the treatment of hard tumors -
PS Example 11; Page -: 61pp; English.
XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 106. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of Intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence.
CC Note: The present sequence is not shown in the specification but is
CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
CC sequence listing (AAE09181).
XX
SQ Sequence 419 AA:
XX
XX Query Match 66.3%; Score 278; DB 22; Length 419;
XX Best local Similarity 100.0%; Pred. No. 1.5e-253;
XX Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 126 TTPVTGASPGGLREQLRLSTECILKGGVLIQRNPOLCYDTILMKDIFKNNQALATLID 185
DB 126 TTPVTGASPGGLREQLRLSTECILKGGVLIQRNPOLCYDTILMKDIFKNNQALATLID 185
QY 186 TNRSRACHPCSPCCCKSRGGESEEDCQSTIRIVCAGGACRGRIPTDCHCQCAAGCT 245
DB 186 TNRSRACHPCSPCCCKSRGGESEEDCQSTIRIVCAGGACRGRIPTDCHCQCAAGCT 245
QY 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDFFESCPNPEGRTGASCVTACPNKLS 305
DB 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDFFESCPNPEGRTGASCVTACPNKLS 305
QY 306 TDVGSCTIVCPHLNQEVTAEADGTQRCCKSPCARGXHSXXXPPRAAVPVXRXQXPAPHP 365
DB 306 TDVGSCTIVCPHLNQEVTAEADGTQRCCKSPCARGXHSXXXPPRAAVPVXRXQXPAPHP 365
QY 366 VLSFLRPSMDXVSATFSLPLAPDPTSVXISPVSVGKXDPDAHVAV 412
DB 366 VLSFLRPSMDXVSATFSLPLAPDPTSVXISPVSVGKXDPDAHVAV 412

RESULT 5
AAE09211
ID AAE09211 standard; Protein; 419 AA.
AC AAE09211;
XX 15-NOV-2001 (first entry)
XX Human p68HER-2 generic protein variant 9.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH Region 1..340
XX /note="Identical to N-terminal region of p185HER-2"
XX FT Domain 341..419
XX /label="ECDIIIA_variant"
XX /note="Extracellular domain IIA variant"
XX

```



```

PF 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11; Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX sequence listing (AAE09181).
XX
XX Sequence 419 AA:
XX
XX Query Match 62.3%; Score 261; DB 22; Length 419;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-237;
XX Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 126 TTPVTGASPGELRELDRLTECLKGGVLIQRPOLCYDTIIMKDFHKNNOIALTLID 185
XX |||||||
XX DB 126 TTPVTGASPGELRELDRLTECLKGGVLIQRPOLCYDTIIMKDFHKNNOIALTLID 185
XX
XX QY 186 TNSRACHPCSPCCCKSGESSEDCQSLTRIVCAGGACRCKGRLPTDCCHEQCAAGCT 245
XX |||||||
XX DB 186 TNSRACHPCSPCCCKSGESSEDCQSLTRIVCAGGACRCKGRLPTDCCHEQCAAGCT 245
XX
XX QY 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDFESCNPNEGRTTEGASCVTACPYKLS 305
XX |||||||
XX DB 246 GPKHSDCLACLFHNSGICELHCPALVTYNTDFESCNPNEGRTTEGASCVTACPYKLS 305
XX
XX QY 306 TDVGSCTLVCPLEHNOEVTADGTORCEKCSKPCARGHSHXXPPRAVPVXXQXPXPAHP 365
XX |||||||
XX DB 306 TDVGSCTLVCPLEHNOEVTADGTORCEKCSKPCARGHSHXXPPRAVPVXXQXPXPAHP 365
XX
XX QY 366 VLSFLRPSMDXVSAPYSLSPLAPLDPTSV 393
XX |||||||
XX DB 366 VLSFLRPSMDXVSAPYSLSPLAPLDPTSV 393
XX
XX RESULT 7
XX AAE09209
XX ID AAE09209 standard; Protein; 419 AA.
XX
XX AC AAE09209;
XX
XX XX 15-NOV-2001 (first entry)
XX
XX DE Human p68HER-2 generic protein variant 7.
XX
XX XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.

```

```

XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Region 1..340 "identical to N-terminal region of p185HER-2"
XX FT Domain /note="Identical to N-terminal region of p185HER-2"
XX FT /label="ECDIIIA-variant"
XX FT /note="Extracellular domain IIRa variant"
XX
XX FT Misc-difference 124 /note="Represented as Agn in the parent sequence shown
XX FT in the specification"
XX FT Misc-difference 125 /note="Represented as Agn in the parent sequence shown
XX FT in the specification"
XX
XX FT Misc-difference 342 /note="Represented as Agn in the parent sequence shown
XX FT in the specification"
XX FT Misc-difference 345 /label="Unknown"
XX FT Misc-difference 346 /label="Unknown"
XX FT Misc-difference 346 /label="Unknown"
XX FT Misc-difference 356 /label="Unknown"
XX FT Misc-difference 358 /label="Unknown"
XX FT Misc-difference 358 /label="Unknown"
XX FT Misc-difference 361 /label="Unknown"
XX FT Misc-difference 376 /label="Unknown"
XX FT /note="p68HER-2 generic sequence (AAE09181) Xaa
XX FT substituted with Ile"
XX
XX FT Misc-difference 394 /label="Unknown"
XX FT Misc-difference 404 /label="Unknown"
XX FT Misc-difference 413 /label="Unknown"
XX
XX PN WO200161356-A1.
XX
XX PD 23-AUG-2001.
XX
XX PF 16-FEB-2001; 2001WO-US05327.
XX
XX PR 16-FEB-2000; 2000US-0506079.
XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI Clinton G, Henner WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 11; Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX sequence listing (AAE09181).

```


ID AAE09207 standard; protein; 419 AA.
 XX AAE09207;
 AC
 XX
 XX 15-NOV-2001 (first entry)
 DT
 XX
 XX Human p68HER-2 generic protein variant 5.
 DE
 XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytosolic; gene therapy;
 KM p68HER-2; ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Region 1..340
 FT /note="identical to N-terminal region of p185HER-2"
 FT Domain 341..419
 FT /label="ECDIIIA variant"
 FT /note="Extracellular domain IITA variant"
 FT Misc-difference 124
 FT /note="Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 125
 FT /note="Represented as Agn in the parent sequence shown
 FT in the specification"
 FT Misc-difference 342
 FT /label="Unknown
 FT Misc-difference 345
 FT /label="Unknown
 FT Misc-difference 346
 FT /label="Unknown
 FT Misc-difference 356
 FT /label="Unknown
 FT Misc-difference 358
 FT /note="p68HER-2 generic sequence (AAE09181) Xaa
 FT substituted with Leu"
 FT Misc-difference 361
 FT /label="Unknown
 FT Misc-difference 376
 FT /label="Unknown
 FT Misc-difference 394
 FT /label="Unknown
 FT Misc-difference 404
 FT /label="Unknown
 FT Misc-difference 413
 FT /label="Unknown
 XX
 XX PN W0200161356-A1.
 XX
 PD 23-AUG-2001.
 XX
 XX 16-FEB-2001; 2001MO-US05327.
 PE
 XX
 XX 16-FEB-2000; 2000US-0506079.
 PR
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 PA
 XX
 XX Clinton G, Hennen WD, Evans A;
 PI
 XX
 XX WPI; 2001-529934/58.
 DR
 XX
 XX New polypeptide, which binds to the extracellular domain of HER-2 for
 PT the treatment of hard tumors -
 XX
 XX
 PS Example 11; Page -: 61pp; English.
 XX
 CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁷8. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of intron 8. The translation product

CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDIIIA variant sequence.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
 CC sequence listing (AAE09181).
 XX
 SQ Sequence 419 AA:
 Query Match 54.4%; Score 228; DB 22; Length 419;
 Best Local Similarity 100.0%; Pred. No. 2,1e-206;
 Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 126 TTPYTGASPGGLBELRLSLTECLKGVLQIRNPQLCYDTIMKDIFFKNNQALATLID 185
 DB 126 TTPYTGASPGGLBELRLSLTECLKGVLQIRNPQLCYDTIMKDIFFKNNQALATLID 185
 QY 186 TNSRACHPCSPCKGRCMGESSEDCQSLTRIVCAGGACARCKGPLPTDCCHQCAAGCT 245
 DB 186 TNSRACHPCSPCKGRCMGESSEDCQSLTRIVCAGGACARCKGPLPTDCCHQCAAGCT 245
 QY 246 GPKHSDCLAFHNHSGICELHCPALVTYNTDFESCPNDEGRYTTGASCVTACPYNKLS 305
 DB 246 GPKHSDCLAFHNHSGICELHCPALVTYNTDFESCPNDEGRYTTGASCVTACPYNKLS 305
 QY 306 TDVGSCTLVCPRLNQETADGTQRCBKSPCARGXHSXPPRAVPVPXR 357
 DB 306 TDVGSCTLVCPRLNQETADGTQRCBKSPCARGXHSXPPRAVPVPXR 357
 RESULT 10
 ID AAE09183 standard; protein; 419 AA.
 XX AAE09183;
 AC
 XX
 XX 15-NOV-2001 (first entry)
 DT
 XX
 XX Human p68HER-2 generic sequence #2.
 DE
 XX
 XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytosolic; gene therapy;
 KM p68HER-2; ECDIIIA.
 XX
 XX OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Region 1..340
 FT /note="identical to N-terminal region of p185HER-2"
 FT Domain 341..419
 FT /label="ECDIIIA
 FT /note="Extracellular domain IITA"
 FT Misc-difference 124
 FT /note="Represented as Agn in the sequence shown in
 FT the specification"
 FT Misc-difference 125
 FT /note="Represented as Agn in the sequence shown in
 FT the specification"
 FT Misc-difference 342
 FT /label="Unknown
 FT /note="Encoded by WCC"
 FT Misc-difference 345
 FT /label="Unknown
 FT /note="Encoded by CYG"
 FT Misc-difference 346
 FT /label="Unknown
 FT /note="Encoded by CYC"
 FT Misc-difference 356

```

FT      /label- Unknown
FT      /note- "Encoded by CWG"
FT      Misc-difference 357
FT      /label- Unknown
FT      /note- "Encoded by YGC"
FT      Misc-difference 358
FT      /label- Unknown
FT      /note- "Encoded by ATR"
FT      Misc-difference 361
FT      /label- Unknown
FT      /note- "Encoded by GNC"
FT      Misc-difference 371
FT      /label- Unknown
FT      /note- "Encoded by AKA"
FT      Misc-difference 376
FT      /label- Unknown
FT      /note- "Encoded by MTA"
FT      Misc-difference 389
FT      /note- "Encoded by AGC"
FT      Misc-difference 394
FT      /label- Unknown
FT      /note- "Encoded by CST"
FT      Misc-difference 404
FT      /label- Unknown
FT      /note- "Encoded by CYG"
FT      Misc-difference 413
FT      /label- Unknown
FT      /note- "Encoded by VAC"
XX      WO200161356-A1.
XX      23-AUG-2001.
XX      16-FEB-2001; 2001WO-US05327.
XX      16-FEB-2000; 2000US-0506079.
XX      (UYOR-) UNIV OREGON HEALTH SCI.
XX      Clinton G, Henner WD, Evans A;
XX      WPI, 2001-529934/58.
XX      N-PSDB; AAD15852.
XX      New polypeptide, which binds to the extracellular domain of HER-2 for
XX      the treatment of hard tumors -
XX      Claim 8; Page 57-58; 61pp; English.
XX      The invention relates to novel HER-2 (herstatin-2) antagonist
XX      particularly a polypeptide that binds to the extracellular domain (ECD)
XX      of HER-2 at a site that is different from the binding site of humanised
XX      antibody, Herceptin, at an affinity of at least 108. The present
XX      invention is based upon the initial discovery of an alternative HER-2
XX      mRNA transcript with 274 bp insert of intron 8. The translation product
XX      of the alternative transcript is a truncated HER-2 protein designated
XX      p68HER-2 which lacks the transmembrane and intracellular domains of
XX      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIa.
XX      The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
XX      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      nucleic acids encoding these are useful to treat, diagnose and identify
XX      solid tumours. The present sequence is human p68HER-2 protein
XX      containing ECDIIa generic sequence.
XX      Sequence 419 AA:
XX
XX      Query Match 54.2%; Score 227; DB 22; Length 419;
XX      Best Local Similarity 100.0%; Pred. No. 1.8e-205;
XX      Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      186 TNRSRACHPCSPCKSGKSGRWGESSSEDGSLTFTVCAGGCARCKGPLPTDCHEGCAAGCT 245
OY      186 TNRSRACHPCSPCKSGKSGRWGESSSEDGSLTFTVCAGGCARCKGPLPTDCHEGCAAGCT 245
DB      246 GPKHSDCLACLFHNSHGICELHCPALVTYNTDFESCPNPEGRYTFGASCVTACPYNKLIS 305
DB      246 GPKHSDCLACLFHNSHGICELHCPALVTYNTDFESCPNPEGRYTFGASCVTACPYNKLIS 305
OY      306 TDVGSCTLVCPILHNOEVTAEADGTORCEKSRPCAGXHSXPRPAVYP 355
OY      306 TDVGSCTLVCPILHNOEVTAEADGTORCEKSRPCAGXHSXPRPAVYP 355
DB      306 TDVGSCTLVCPILHNOEVTAEADGTORCEKSRPCAGXHSXPRPAVYP 355

RESULT 11
AAE09200
ID      AAE09200 standard; Protein: 419 AA.
AC      AAE09200;
XX      15-NOV-2001 (first entry)
DE      Human p68HER-2 generic protein variant (Arg357Cys).
XX      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX      solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX      p68HER-2; ECDIIa; variant.
XX      Homo sapiens.
XX      Key Location/Qualifiers
XX      Region 1..340
XX      Domain 341..419
XX      /label- ECDIIa,variant
XX      /note- "Extracellular domain IIIa variant"
XX      Misc-difference 124
XX      /note- "Represented as Agn in the parent sequence shown
XX      in the specification"
XX      Misc-difference 125
XX      /note- "Represented as Agn in the parent sequence shown
XX      in the specification"
XX      Misc-difference 342
XX      /label- Unknown
XX      Misc-difference 345
XX      /label- Unknown
XX      Misc-difference 346
XX      /label- Unknown
XX      Misc-difference 356
XX      /label- Unknown
XX      Misc-difference 357
XX      /label- Unknown
XX      /note- "p68HER-2 generic sequence (AAE09181) Arg
XX      substituted with Cys"
XX      Misc-difference 358
XX      /label- Unknown
XX      Misc-difference 361
XX      /label- Unknown
XX      Misc-difference 376
XX      /label- Unknown
XX      Misc-difference 394
XX      /label- Unknown
XX      Misc-difference 404
XX      /label- Unknown
XX      Misc-difference 413
XX      /label- Unknown
XX      WO200161356-A1.
XX      23-AUG-2001.
XX      16-FEB-2001; 2001WO-US05327.
XX      16-FEB-2000; 2000US-0506079.

```

```

XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G., Henner WD, Evans A;
XX PI MPI: 2001-529934/58.
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for
XX PT the treatment of hard tumors -
XX PS Example 12; Page -: 61pp; English.
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence.
CC Note: The present sequence is not shown in the specification but is
CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
CC sequence listing (AAE09181).
XX SQ Sequence 419 AA:

Query Match 54.2%; Score 227; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.8e-205;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 TTPVTASPGELRELRLSTELKGGVLIQRNPOLCYDTIIMKDIFFHNNQALATLID 185
DB 126 TTPVTASPGELRELRLSTELKGGVLIQRNPOLCYDTIIMKDIFFHNNQALATLID 185
QY 186 TNRSRACHPCSPCKSGRCGESESDQSLTRIVCAGGCGARCKGPIPTDCHEGCAAGCT 245
DB 186 TNRSRACHPCSPCKSGRCGESESDQSLTRIVCAGGCGARCKGPIPTDCHEGCAAGCT 245
QY 246 GPRHSDCLACLFHNSGICELHCPALVTYNTDTFESCPNDEGRYRTGASCVTACPYNKLS 305
DB 246 GPRHSDCLACLFHNSGICELHCPALVTYNTDTFESCPNDEGRYRTGASCVTACPYNKLS 305
QY 306 TDVGSGCTIVCPRLHNOEVTADGTQRCCEKSKPCARGXHSXXPRPAVPVP 355
DB 306 TDVGSGCTIVCPRLHNOEVTADGTQRCCEKSKPCARGXHSXXPRPAVPVP 355

RESULT 12
AAE09206
ID AAE09206 standard; Protein; 419 AA.
XX
XX AAE09206;
AC
XX
XX 15-NOV-2001 (first entry)
DT
XX
XX Human p68HER-2 generic protein variant 4.
DE
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 1..340
FT /note="Identical to N-terminal region of p185HER-2"
FT Domain 341..419

```

```

FT FT /label= ECDIIIA-variant
FT FT /note="Extracellular domain IIA variant"
FT FT Misc-difference 124
FT FT /note="Represented as Agn in the parent sequence shown
FT FT in the specification"
FT FT Misc-difference 125
FT FT /note="Represented as Agn in the parent sequence shown
FT FT in the specification"
FT FT Misc-difference 342
FT FT /label= Unknown
FT FT Misc-difference 345
FT FT /label= Unknown
FT FT Misc-difference 346
FT FT /label= Unknown
FT FT Misc-difference 356
FT FT /note="p68HER-2 generic sequence (AAE09181) Xaa
FT FT substituted with Gln"
FT FT Misc-difference 358
FT FT /label= Unknown
FT FT Misc-difference 361
FT FT /label= Unknown
FT FT Misc-difference 376
FT FT /label= Unknown
FT FT Misc-difference 394
FT FT /label= Unknown
FT FT Misc-difference 404
FT FT /label= Unknown
FT FT Misc-difference 413
FT FT /label= Unknown
XX XX WO200161356-A1.
XX XX
XX PD 23-AUG-2001.
XX XX
XX PF 16-FEB-2001; 2001WO-US05327.
XX XX
XX PR 16-FEB-2000; 2000US-0506079.
XX XX
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX PI Clinton G., Henner WD, Evans A;
XX PI MPI: 2001-529934/58.
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for
XX PT the treatment of hard tumors -
XX PS Example 11; Page -: 61pp; English.
CC The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10-8. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence.
CC Note: The present sequence is not shown in the specification but is
CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
CC sequence listing (AAE09181).
XX SQ Sequence 419 AA:

Query Match 54.2%; Score 227; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.8e-205;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

OY 126 TTPVTGASPGGIREQLNSLRECLKGVLIQRNPOLCYQDTILMKDIFHKNNQALATLTD 185
XX |||||||
XX 126 TTPVTGASPGGIREQLNSLRECLKGVLIQRNPOLCYQDTILMKDIFHKNNQALATLTD 185
Db |||||||
OY 186 TNRSRACHPCSPCCGSKGSRGWGSSSDCOSLTFTVAGAGCARCKGRLPTDCCHEGCAAGCT 245
XX |||||||
XX 186 TNRSRACHPCSPCCGSKGSRGWGSSSDCOSLTFTVAGAGCARCKGRLPTDCCHEGCAAGCT 245
Db TNRSRACHPCSPCCGSKGSRGWGSSSDCOSLTFTVAGAGCARCKGRLPTDCCHEGCAAGCT 245
OY 246 GPKHSDCLACLFHNSGICELHCPALVYNTDTFESCPNPEGRTYFGASCYACPYNKLS 305
XX |||||||
XX 246 GPKHSDCLACLFHNSGICELHCPALVYNTDTFESCPNPEGRTYFGASCYACPYNKLS 305
Db GPKHSDCLACLFHNSGICELHCPALVYNTDTFESCPNPEGRTYFGASCYACPYNKLS 305
OY 306 TDVGSCTLVCPILHNOEVTAEADGTORCEKSKPCARGXHSXPRPAVPVP 355
XX |||||||
XX 306 TDVGSCTLVCPILHNOEVTAEADGTORCEKSKPCARGXHSXPRPAVPVP 355
Db TDVGSCTLVCPILHNOEVTAEADGTORCEKSKPCARGXHSXPRPAVPVP 355

RESULT 13
AAE09214
ID AAE09214 standard; Protein: 419 AA.
XX
AC AAE09214;
XX
DE 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 13.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumor; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX
FH Key 1.340 Location/Qualifiers
FT Region /note= "Identical to N-terminal region of p185HER-2"
FT Domain 341..419 /label= "ECDIIIA_variant"
FT /note= "Extracellular domain IIIa"
FT Misc-difference 124 /note= "Represented as Agn in the sequence shown in
FT /note= "the specification"
FT Misc-difference 125 /note= "Represented as Agn in the sequence shown in
FT /note= "the specification"
FT Misc-difference 342 /label= "Unknown"
FT Misc-difference 343 /label= "Unknown"
FT Misc-difference 345 /label= "Unknown"
FT Misc-difference 346 /label= "Unknown"
FT Misc-difference 346 /label= "Unknown"
FT Misc-difference 356 /label= "Unknown"
FT Misc-difference 357 /label= "Unknown"
FT /note= "p68HER-2 generic sequence (AAE09183) Xaa
FT /note= "substituted with Cys"
FT Misc-difference 358 /label= "Unknown"
FT Misc-difference 361 /label= "Unknown"
FT Misc-difference 371 /label= "Unknown"
FT Misc-difference 376 /label= "Unknown"
FT Misc-difference 394 /label= "Unknown"
FT Misc-difference 394 /label= "Unknown"
FT Misc-difference 404 /label= "Unknown"
FT Misc-difference 413 /label= "Unknown"
FT Misc-difference 413 /label= "Unknown"
XX
XX WO200161356-A1.
XX

```

```

PD 23-AUG-2001.
XX
XX 16-FEB-2001; 2001WO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Henner WD, Evans A;
XX
XX WPI: 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 12: Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 1078. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is
XX derived from p68HER-2 generic sequence (SEQ ID NO:13) shown in the
XX sequence listing (AAE09183).
XX
XX Sequence 419 AA:
XX
XX Query Match 54.2%; Score 227; DB 22; Length 419;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-205;
XX Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 126 TTPVTGASPGGIREQLNSLRECLKGVLIQRNPOLCYQDTILMKDIFHKNNQALATLTD 185
XX |||||||
XX 126 TTPVTGASPGGIREQLNSLRECLKGVLIQRNPOLCYQDTILMKDIFHKNNQALATLTD 185
Db |||||||
OY 186 TNRSRACHPCSPCCGSKGSRGWGSSSDCOSLTFTVAGAGCARCKGRLPTDCCHEGCAAGCT 245
XX |||||||
XX 186 TNRSRACHPCSPCCGSKGSRGWGSSSDCOSLTFTVAGAGCARCKGRLPTDCCHEGCAAGCT 245
Db TNRSRACHPCSPCCGSKGSRGWGSSSDCOSLTFTVAGAGCARCKGRLPTDCCHEGCAAGCT 245
OY 246 GPKHSDCLACLFHNSGICELHCPALVYNTDTFESCPNPEGRTYFGASCYACPYNKLS 305
XX |||||||
XX 246 GPKHSDCLACLFHNSGICELHCPALVYNTDTFESCPNPEGRTYFGASCYACPYNKLS 305
Db GPKHSDCLACLFHNSGICELHCPALVYNTDTFESCPNPEGRTYFGASCYACPYNKLS 305
OY 306 TDVGSCTLVCPILHNOEVTAEADGTORCEKSKPCARGXHSXPRPAVPVP 355
XX |||||||
XX 306 TDVGSCTLVCPILHNOEVTAEADGTORCEKSKPCARGXHSXPRPAVPVP 355
Db TDVGSCTLVCPILHNOEVTAEADGTORCEKSKPCARGXHSXPRPAVPVP 355

RESULT 14
AAE09215
ID AAE09215 standard; Protein: 419 AA.
XX
AC AAE09215;
XX
DE 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 14.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumor; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
OS Homo sapiens.
XX

```

```

XX Key Location/Qualifiers
FH Region 1..340
FT /note="Identical to N-terminal region of p185HER-2"
FT 341..419
FT Domain /label="ECDI1A variant"
FT /note="Extracellular domain IIIa"
FT Misc-difference 124
FT /note="Represented as Agn in the sequence shown in
FT the specification"
FT Misc-difference 125
FT /note="Represented as Agn in the sequence shown in
FT the specification"
FT Misc-difference 342
FT /label="Unknown"
FT Misc-difference 345
FT /label="Unknown"
FT Misc-difference 346
FT /label="Unknown"
FT Misc-difference 356
FT /label="Unknown"
FT Misc-difference 357
FT /label="Unknown"
FT Misc-difference 358
FT /label="Unknown"
FT Misc-difference 361
FT /label="Unknown"
FT Misc-difference 371
FT /note="p68HER-2 generic sequence (AAE09183) Xaa
FT substituted with Ile"
FT Misc-difference 376
FT /label="Unknown"
FT Misc-difference 394
FT /label="Unknown"
FT Misc-difference 404
FT /label="Unknown"
FT Misc-difference 413
FT /label="Unknown"
FT Misc-difference 413
FT /label="Unknown"
PN WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001, 2001WO-US05327.
XX
XX 16-FEB-2000, 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Hennen WD, Evans A;
XX
XX WPI; 2001-529934/58.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
XX Example 12; Page -: 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 106. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX nucleic acids encoding these are useful to treat, diagnose and identify
XX solid tumours. The present sequence is human p68HER-2 generic protein
XX containing ECDIIIA variant sequence.
XX Note: The present sequence is not shown in the specification but is

```

```

CC derived from p68HER-2 generic sequence (SEQ ID NO:13) shown in the
CC sequence listing (AAE09183).
XX
XX Sequence 419 AA;
SQ
Query Match 54.2%; Score 227; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.8e-205;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 126 TTPYTGASPGALRELORSLTECLKGVLTQORNPOLCYOPTILMKDIFHNKNOIALTLD 185
DB 126 TTPYTGASPGALRELORSLTECLKGVLTQORNPOLCYOPTILMKDIFHNKNOIALTLD 185
QY 186 TNRSRACHPCSPCKGSRGWGSESDCQSLRTVACGACARCKGPLEPTDCCHQCAAGCT 245
DB 186 TNRSRACHPCSPCKGSRGWGSESDCQSLRTVACGACARCKGPLEPTDCCHQCAAGCT 245
QY 246 GPKHSDCLACLFHNHSGICELHCPALVTYNTDFFESCPNPEGRTFASCVTACPYNKLS 305
DB 246 GPKHSDCLACLFHNHSGICELHCPALVTYNTDFFESCPNPEGRTFASCVTACPYNKLS 305
QY 306 TDVGSCTLVCPNLHNOEYTAEDGTORCEKSKPCARGXHSXPPRAVVP 355
DB 306 TDVGSCTLVCPNLHNOEYTAEDGTORCEKSKPCARGXHSXPPRAVVP 355
RESULT 15
AAE09216
ID AAE09216 standard; Protein; 419 AA.
XX
XX AAE09216;
XX
XX 15-NOV-2001 (first entry)
XX
XX Human p68HER-2 generic protein variant 15.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX p68HER-2; ECDIIIA; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 1..340
FT /note="Identical to N-terminal region of p185HER-2"
FT 341..419
FT Domain /label="ECDI1A variant"
FT /note="Extracellular domain IIIa"
FT Misc-difference 124
FT /note="Represented as Agn in the sequence shown in
FT the specification"
FT Misc-difference 125
FT /note="Represented as Agn in the sequence shown in
FT the specification"
FT Misc-difference 342
FT /label="Unknown"
FT Misc-difference 345
FT /label="Unknown"
FT Misc-difference 346
FT /label="Unknown"
FT Misc-difference 356
FT /label="Unknown"
FT Misc-difference 357
FT /label="Unknown"
FT Misc-difference 358
FT /label="Unknown"
FT Misc-difference 361
FT /label="Unknown"
FT Misc-difference 371
FT /label="Unknown"
FT Misc-difference 376
FT /label="Unknown"
FT Misc-difference 394

```

```

FT      /label= Unknown
FT      Misc-difference 404
FT      /label= unknown
FT      Misc-difference 413
FT      /note= "p68HER-2 generic sequence (AAE09183) Xaa
FT      substituted with Asn"
XX      WO200161356-A1.
XX      23-AUG-2001.
XX      PD
XX      PF 16-FEB-2001; 2001WO-US05327.
XX      PR 16-FEB-2000; 2000US-0506079.
XX      PA (UYOR-) UNIV OREGON HEALTH SCI.
XX      PI Clinton G, Henner WD, Evans A;
XX      DR WPI; 2001-529934/58.
XX      PT
XX      PS New polypeptide, which binds to the extracellular domain of HER-2 for
XX      the treatment of hard tumors -
XX      Example 12; Page -: 61pp; English.
XX      CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX      CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX      CC of HER-2 at a site that is different from the binding site of humanised
XX      CC antibody, Herceptin, at an affinity of at least 108. The present
XX      CC invention is based upon the initial discovery of an alternative HER-2
XX      CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX      CC of the alternative transcript is a truncated HER-2 protein designated
XX      CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX      CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIa.
XX      CC The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
XX      CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      CC nucleic acids encoding these are useful to treat, diagnose and identify
XX      CC solid tumours. The present sequence is human p68HER-2 generic protein
XX      CC containing ECDIIa variant sequence.
XX      CC Note: The present sequence is not shown in the specification but is
XX      CC derived from p68HER-2 generic sequence (SEQ ID NO:13) shown in the
XX      CC sequence listing (AAE09183).
XX      SQ Sequence 419 AA;

Query Match 54.2%; Score 227; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.8e-205;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 TTPVTGASPGGLREQLRLTECLKGVLIQRNPQLCYODTILMKDIFHKNNQLALTLID 185
DB 126 TTPVTGASPGGLREQLRLTECLKGVLIQRNPQLCYODTILMKDIFHKNNQLALTLID 185
QY 186 TNRSRACHPCSPCKSGKSWESSSEDCSLTRTYVAGGACARCKGRLPTDCHEQCAGCT 245
DB 186 TNRSRACHPCSPCKSGKSWESSSEDCSLTRTYVAGGACARCKGRLPTDCHEQCAGCT 245
QY 246 GPKHSDCLACHFNHNSGICELHCPALVYNTDFESCPNPEGRTYFGASCVTACPYNKLIS 305
DB 246 GPKHSDCLACHFNHNSGICELHCPALVYNTDFESCPNPEGRTYFGASCVTACPYNKLIS 305
QY 306 TDVGSCTLVCPHLNQEVTAEDGTORCEKCRKPCARGXSHXXPRPAAVPVP 355
DB 306 TDVGSCTLVCPHLNQEVTAEDGTORCEKCRKPCARGXSHXXPRPAAVPVP 355

```

Search completed: January 14, 2003, 17:12:55
 Job time : 54.8474 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:10:43 ; Search time 8.7249 Seconds
(without alignments)
1865.663 Million cell updates/sec

Title: US-09-506-079E-1

Perfect score: 1 GMSXPRPAAPVPRXRP.....VGRGDPDAHVXLSRREG 79

Sequence:

Scoring table: OLIGO

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	21.5	419	4 Q9UK79	Q9UK79 homo sapien
2	7	8.9	27	4 Q9HD19	Q9HD19 homo sapien
3	7	8.9	121	16 Q8Y2M8	Q8Y2M8 arabidopsi
4	7	8.9	178	10 Q9FWC7	Q9FWC7 arabidopsi
5	7	8.9	245	4 Q14596	Q14596 homo sapien
6	7	8.9	268	2 Q9EWB9	Q9EWB9 streptomyc
7	7	8.9	382	5 Q9U3E1	Q9U3E1 caenorhabdi
8	7	8.9	389	16 Q8UEF8	Q8UEF8 agrobacteri
9	7	8.9	416	16 Q9KZK5	Q9KZK5 streptomyc
10	7	8.9	436	16 Q67814	Q67814 aquilex aeo
11	7	8.9	469	5 Q8W063	Q8W063 drosophila
12	7	8.9	469	5 Q8WPA9	Q8WPA9 drosophila
13	7	8.9	469	5 Q8WPA9	Q8WPA9 drosophila
14	7	8.9	471	5 Q8W065	Q8W065 drosophila
15	7	8.9	471	5 Q8W064	Q8W064 drosophila
16	7	8.9	471	5 Q8WPC6	Q8WPC6 drosophila

17	7	8.9	471	5 Q8WPC8	Q8WPC8 drosophila
18	7	8.9	471	5 Q8WPA6	Q8WPA6 drosophila
19	7	8.9	473	5 Q8W066	Q8W066 drosophila
20	7	8.9	473	5 Q8WPA9	Q8WPA9 drosophila
21	7	8.9	475	5 Q8W067	Q8W067 drosophila
22	7	8.9	481	10 Q945M4	Q945M4 arabidopsi
23	7	8.9	493	2 Q52799	Q52799 amycolatops
24	7	8.9	520	16 Q66518	Q66518 aquilex aeo
25	7	8.9	595	5 Q18900	Q18900 caenorhabdi
26	7	8.9	635	4 Q969Y5	Q969Y5 homo sapien
27	7	8.9	798	2 Q8VR14	Q8VR14 myxococcus
28	7	8.9	865	10 Q22774	Q22774 arabidopsi
29	7	8.9	879	10 Q41010	Q41010 pisum sativ
30	7	8.9	1119	4 Q75762	Q75762 homo sapien
31	7	8.9	1469	10 Q9LKR1	Q9LKR1 pisum sativ
32	7	8.9	1503	10 Q81283	Q81283 arabidopsi
33	7	8.9	1599	16 Q8R714	Q8R714 thermoaer
34	7	8.9	2240	16 Q9S129	Q9S129 streptomyc
35	7	8.9	19	12 Q84862	Q84862 unidentified
36	7	7.6	19	12 Q84863	Q84863 unidentified
37	7	7.6	25	11 Q9JRY4	Q9JRY4 mus musculu
38	7	7.6	39	5 Q8W0U2	Q8W0U2 plasmodium
39	6	7.6	50	16 Q8VJ18	Q8VJ18 mycobacteri
40	6	7.6	54	12 Q91J10	Q91J10 white spot
41	6	7.6	65	17 Q9YB46	Q9YB46 aeropyrum p
42	6	7.6	74	16 Q9PFA2	Q9PFA2 xylella fas
43	6	7.6	82	2 Q50088	Q50088 mycobacteri
44	6	7.6	82	10 Q8RZY5	Q8RZY5 oryza sativ
45	6	7.6	86	9 Q9MC37	Q9MC37 bacterioph

ALIGNMENTS

RESULT 1

Q9UK79 ID Q9UK79 PRELIMINARY: PRT: 419 AA.
AC Q9UK79:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Herstatin.
GN HER-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99415951; PubMed=10485918;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted
autoinhibitor";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Henner W.D.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF177761; AAD56009.2;
DR InterPro; IPR000494; EGFR_L-domain.
DR Pfam; PF00757; Furin-Like; 1.
DR Pfam; PF01030; Recep_L-domain; 1.
DR SMART; SM00261; FU; 1.
DR SEQUENCE 419 AA; 45472 MW; FECLIB347E2D030C CXC64;

Query Match 21.5%; Score 17; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 VSAFSLPLAPSPSV 53
DB . 377 VSAFSLPLAPSPSV 393

```

RESULT 2
Q9HD19          PRELIMINARY:      PRT:      27 AA.
AC  Q9HD19:
DT  01-MAR-2001 (TReMBLrel. 16, Created)
DT  01-MAR-2001 (TReMBLrel. 16, last sequence update)
DT  01-JUN-2002 (TReMBLrel. 21, last annotation update)
DE  Sodium-dependent multivitamin transporter (Fragment).
GN  SMVT OR SIC5A6.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX  NCBI_TaxID=9606;
RN  11
RP  SEQUENCE FROM N.A.
RC  TISSUE=HEART;
RA  Rubin S.A., Dey S., Chatterjee N., Said H.M.;
RT  "Molecular characterization of the human heat SMVT cDNA.";
RL  Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN  12
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21952373; PubMed=11955628;
RA  Dey S., Subramanian V.S., Chatterjee N.S., Rubin S.A., Said H.M.;
RT  "Characterization of the 5' regulatory region of the human sodium-
RT  dependent multivitamin transporter, hSMVT.";
RL  Biochim. Biophys. Acta 1574:187-192(2002).
DR  EMBL; AF288781; AAG00587.1; -
DR  EMBL; AF442149; AAL84706.1; -
DR  EMBL; AF442150; AAL84707.1; -
FT  NON_TER
SQ  SEQUENCE 27 AA; 2620 MW; 27CDCC38BD55E9D CRC64;

Query Match      8.9%; Score 7; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  46 APLSPTS 52
    |||||
Db  9 APLSPTS 15

RESULT 3
Q8YZW8          PRELIMINARY:      PRT:      121 AA.
AC  Q8YZW8:
DT  01-MAR-2002 (TReMBLrel. 20, Created)
DT  01-MAR-2002 (TReMBLrel. 20, last sequence update)
DT  01-JUN-2002 (TReMBLrel. 21, last annotation update)
DE  Hypothetical Protein A110337.
GN  A110337.
OS  Anabaena sp. (strain PCC 7120).
OC  Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX  NCBI_TaxID=103690;
RN  11
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21595285; PubMed=11759840;
RA  Kaneo T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA  Watanabe A., Iriuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA  Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA  Nakazaki N., Shilpo S., Sugimoto M., Takazawa M., Yamada M.,
RA  Yasuda M., Tabata S.;
RT  "Complete genomic sequence of the filamentous nitrogen-fixing
RT  cyanobacterium Anabaena sp. strain PCC 7120.";
RL  DNA Res. 8:205-213(2001).
DR  EMBL; AP003582; BAB72295.1; -
DR  InterPro; IPR003477; PsmK.
DR  Pfam; PF02452; PsmK; 1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 121 AA; 13266 MW; 8E01290F333B1B7A CRC64;

Query Match      8.9%; Score 7; DB 16; Length 121;

```

```

Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  25 PVLSPFLR 31
    |||||
Db  56 PVLSPFLR 62

RESULT 4
Q9FMC7          PRELIMINARY:      PRT:      178 AA.
AC  Q9FMC7:
DT  01-MAR-2001 (TReMBLrel. 16, Created)
DT  01-MAR-2001 (TReMBLrel. 16, last sequence update)
DT  01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE  Hypothetical 20.3 kDa protein.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX  NCBI_TaxID=3702;
RN  11
RP  SEQUENCE FROM N.A.
RC  STRAIN=COLUMBIA;
RX  MEDLINE=98162728; PubMed=9501997;
RA  Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA  Tabata S.;
RT  "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT  Sequence features of the regions of 1,191,918 bp covered by seventeen
RT  physically assigned P1 clones.";
RL  DNA Res. 4:401-414(1997).
RN  12
RP  SEQUENCE FROM N.A.
RA  Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA  Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA  Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA  Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA  Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA  Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA  Theologis A.;
RT  "Full length cDNA of gene MOK11.15/AT5G04830 (GI:9758458).";
RL  Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN  13
RP  SEQUENCE FROM N.A.
RA  Yamada K., Liu S.X., Pham P.K., Banh J., Banno F., Dale J.M.,
RA  Goldsmith A.D., Jiang P.X., Lee J.M., Ondera C.S., Quach H.L.,
RA  Tang C., Toriumi M., Yamamura Y., Yu G., Yu S., Bowser L.,
RA  Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA  Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koestema E., Lam B.,
RA  Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA  Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA  Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT  "Full length cDNA of gene MOK11.15/AT5G04830 (GI:9758458).";
RL  Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AB008271; BAB08987.1; -
DR  EMBL; AF360335; AAK28632.1; -
DR  EMBL; AY051070; AAK93747.1; -
KW  Hypothetical protein.
SQ  SEQUENCE 178 AA; 20294 MW; 955AC984A4A07FE54 CRC64;

Query Match      8.9%; Score 7; DB 10; Length 178;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  38 SAFYSLP 44
    |||||
Db  66 SAFYSLP 72

RESULT 5
Q14596          PRELIMINARY:      PRT:      245 AA.
AC  Q14596:

```

DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Folate binding protein.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary Gland;
 RA Verma R.S., Elwood P.C.;
 RT "Identification of a homologous cDNA to folate receptor from human
 RT salivary gland."
 RT Submitted (APR-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF000380; AAB81937.1; -
 DR InterPro: IPR004269; Folate_rec.
 DR Pfam: PF03024; Folate_rec.1.
 SQ SEQUENCE 245 AA; 26268 MW; BA36B7DB95FC0E20 CRC64;

Query Match 8.9%; Score 7; DB 4; Length 245;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 APLSPTS 52
 Db 187 APLSPTS 193

RESULT 6
 ID Q9EMB9 PRELIMINARY; PRT; 268 AA.

AC Q9EMB9;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ABC-transporter.
 GN CANR.
 OS Streptomyces griseus.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IMRU 3570;
 RA Campelo A.B., Gill J.A.;
 RT "Cloning and characterization of a gene cluster from Streptomyces
 RT griseus IMRU 3570 involved in candididin production."
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IMRU 3570;
 RA Campelo A.B.;
 RL Thesis (2000), Department of Microbiology, Universidad de Leon.
 DR EMBL: AJ300302; CAC22119.1; -
 DR InterPro: IPR000412; ABCtransprt2.
 DR Pfam: PF01061; ABC2_membrane.1.
 DR PRINTS: PR00164; ABC2TRANSPO.
 DR PROSITE: PS00890; ABC2_MEMBRANE; UNKNOWN.1.
 SQ SEQUENCE 268 AA; 27945 MW; C621E7IDE2EFF7B4 CRC64;

Query Match 8.9%; Score 7; DB 2; Length 268;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 PRPAVP 13
 Db 2 PRPAVP 8

RESULT 7
 ID Q9U3E1 PRELIMINARY; PRT; 382 AA.

AC Q9U3E1;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE F58A4.7b protein.
 GN F58A4.7B.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Berkis M.;
 RL Submitted (APR-1993) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RT Science 282:2012-2018(1998).
 DR EMBL: Z22179; CA80170.1; -
 DR HSSP: P36956; IAM9.
 DR InterPro: IPR001092; HLH_basic.
 DR Pfam: PF00010; HLH.1.
 DR SMART: SM00353; HLH.1.
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN.1.
 SQ SEQUENCE 382 AA; 40763 MW; 510BEF225B073804 CRC64;

Query Match 8.9%; Score 7; DB 5; Length 382;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 APLSPTS 52
 Db 48 APLSPTS 54

RESULT 8

ID Q8UEF8 PRELIMINARY; PRT; 389 AA.

AC Q8UEF8;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical protein Atu1800.
 GN Atu1800 OR AGK_C_3311.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Rhizobium.
 OC NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Boye D. Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Neeter E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58."
 RT Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
 RA Urolo B., Goldman B.S., Cao Y., Askenazi M., Hailing C., Mullin L.,
 RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,

RA Mollam C., Allinger M., Doughty D., Scott C., Iappas C., Markelz B.,
 RA Pliagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.,
 RT "genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL: AEO09135; AAL42798.1; -.
 DR EMBL: AEO08101; AAK87570.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 389 AA; 40230 MW; DED360C24024E93B CRC64;

Query Match 8.9%; Score 7; DB 16; Length 389;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PAAVVP 15
 Db 160 PAAVVP 166

RESULT 9
 Q9KZK5 PRELIMINARY; PRT; 416 AA.
 AC Q9KZK5;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative secreted protein.
 GN SCO3040 OR SCE34.21C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA MEDLINE=97000351; PubMed=843436;
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)";
 RL Nature 417:141-147(2002).
 DR EMBL: AJ353862; CAB88924.1; -.
 DR HSSP: P00806; ILBA.
 SQ SEQUENCE 416 AA; 43715 MW; 92CEAE6017968445 CRC64;

Query Match 8.9%; Score 7; DB 16; Length 416;
 Best Local Similarity 100.0%; Pred. No. 43;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 42 SLPLAPL 48
 Db 52 SLPLAPL 58

RESULT 10
 O67814 PRELIMINARY; PRT; 436 AA.
 ID O67814;
 AC O67814;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Protoporphyrinogen oxidase.
 GN HEMG OR AO_2015.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxId=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Shead W.A., Keller M., Aulay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 DR EMBL: AEO00768; AAC07778.1; -.
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR000205; NAD_binding.
 DR InterPro: IPR004572; Proto_IX_ox.
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR TRIGRAMS: TRGR00562; proto_IX_ox; 1.
 KW Complete proteome.
 SQ SEQUENCE 436 AA; 48987 MW; ED5F2B1BCD1DEEF7 CRC64;

Query Match 8.9%; Score 7; DB 16; Length 436;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 VLSFLRP 32
 Db 428 VLSFLRP 434

RESULT 11
 Q8W063 PRELIMINARY; PRT; 469 AA.
 ID Q8W063;
 AC Q8W063;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Zeste protein (Fragment).
 GN ZESTE.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Phlebotominae; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyroidae; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K13;
 RA Zangerl B.;
 RT "Evidence for selection in a natural population of Drosophila
 melanogaster.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ399793; CAC81400.1; -.
 FT NON_TER 1 1
 FT 469 469
 SQ SEQUENCE 469 AA; 51012 MW; 5F3399F575C6CE24 CRC64;

Query Match 8.9%; Score 7; DB 5; Length 469;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVP 15
|||||||
DB 362 PAAVPVP 368

RESULT 12
O8WP54 PRELIMINARY; PRT; 469 AA.
AC O8WP54; PRT; 469 AA.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Zeste protein (Fragment).
GN ZESTE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K2, K19, AND K22;
RA Zangerl B.;
RT "Evidence for selection in a natural population of Drosophila melanogaster."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ399788; CAC81395.1; -
DR EMBL; AJ399795; CAC81402.1; -
DR EMBL; AJ399798; CAC81405.1; -
FT NON_TER 1 1
FT NON_TER 469 469
SQ SEQUENCE 469 AA; 50996 MW; 474BF9F575D4DC24 CRC64;

Query Match 8.9%; Score 7; DB 5; Length 469;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVP 15
|||||||
DB 362 PAAVPVP 368

RESULT 13
O8WP54 PRELIMINARY; PRT; 469 AA.
AC O8WP54; PRT; 469 AA.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Zeste protein (Fragment).
GN ZESTE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ND6, ND11, ND13, AND ND19;
RA Zangerl B.;
RT "Evidence for selection in a natural population of Drosophila melanogaster."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ399747; CAC81411.1; -
DR EMBL; AJ399748; CAC81412.1; -
DR EMBL; AJ399749; CAC81413.1; -
DR EMBL; AJ399751; CAC81415.1; -
DR EMBL; AJ399756; CAC81421.1; -
FT NON_TER 1 1

FT NON_TER 469 469
SQ SEQUENCE 469 AA; 51012 MW; D8C09D3CCF35AF36 CRC64;

Query Match 8.9%; Score 7; DB 5; Length 469;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVP 15
|||||||
DB 362 PAAVPVP 368

RESULT 14
O8WO65 PRELIMINARY; PRT; 471 AA.
AC O8WO65; PRT; 471 AA.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Zeste protein (Fragment).
GN ZESTE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K8;
RA Zangerl B.;
RT "Evidence for selection in a natural population of Drosophila melanogaster."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ399789; CAC81396.1; -
FT NON_TER 1 1
FT NON_TER 471 471
SQ SEQUENCE 471 AA; 51194 MW; 07E4221753E96B20 CRC64;

Query Match 8.9%; Score 7; DB 5; Length 471;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVPVP 15
|||||||
DB 364 PAAVPVP 370

RESULT 15
O8WO64 PRELIMINARY; PRT; 471 AA.
AC O8WO64; PRT; 471 AA.
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Zeste protein (Fragment).
GN ZESTE.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K11;
RA Zangerl B.;
RT "Evidence for selection in a natural population of Drosophila melanogaster."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ399791; CAC81398.1; -
FT NON_TER 1 1
FT NON_TER 471 471
SQ SEQUENCE 471 AA; 51199 MW; 075DE4C342F9CE4 CRC64;

Query Match 8.9%; Score 7; DB 5; Length 471;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 PAAVPVP 15
| | | | | | |
Db 364 PAAVPVP 370

Search completed: January 14, 2003, 17:14:22
Job time : 11.7249 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:14:40 ; Search time 2.69679 Seconds

(without alignments)
568.337 Million cell updates/sec

Title: US-09-506-079E-1

Perfect score: 1 GMSXXPRPAVFPKXKP.....VGRGPDHVAVYLSRYEG 79

Sequence: OLIGO

Scoring table: Gapop 60.0 , Gapext 60.0

Searched: 118974 seqs, 19401057 residues

Word size : 0

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	8.9	61	10	US-09-864-761-44740
2	7	8.9	215	9	US-10-001-876-210
3	7	7.6	38	10	US-09-864-761-41953
4	6	7.6	68	10	US-09-764-887-226
5	6	7.6	86	10	US-09-764-869-1118
6	6	7.6	168	10	US-09-925-301-1470
7	6	7.6	238	12	US-10-024-579-16
8	6	7.6	257	12	US-10-024-579-14
9	6	7.6	264	12	US-10-024-579-12
10	6	7.6	267	9	US-09-808-602-23
11	6	7.6	267	9	US-09-808-602-25
12	6	7.6	283	12	US-10-024-579-10
13	6	7.6	290	9	US-10-041-006A-7
14	6	7.6	290	12	US-10-028-072-222
15	6	7.6	290	12	US-10-040-655-7
16	6	7.6	298	10	US-09-764-853-552
17	6	7.6	299	9	US-09-808-602-63
18	6	7.6	323	10	US-09-764-864-1265
19	6	7.6	349	10	US-09-766-366-4

20	6	7.6	382	9	US-09-941-947A-30	Sequence 30, Appl
21	6	7.6	398	10	US-09-741-669-430	Sequence 430, App
22	6	7.6	404	10	US-09-764-864-1244	Sequence 1244, Ap
23	6	7.6	408	9	US-10-027-806-74	Sequence 74, Appl
24	6	7.6	408	9	US-10-034-623-74	Sequence 74, Appl
25	6	7.6	418	10	US-09-815-242-12003	Sequence 12003, A
26	6	7.6	450	10	US-09-764-864-804	Sequence 804, App
27	6	7.6	685	9	US-10-029-180-82	Sequence 82, Appl
28	6	7.6	720	10	US-09-919-497-83	Sequence 83, Appl
29	6	7.6	784	10	US-09-770-595A-3	Sequence 3, Appl1
30	6	7.6	912	9	US-10-047-542-75	Sequence 75, Appl1
31	6	7.6	1213	10	US-09-815-242-5358	Sequence 5358, Ap
32	6	7.6	1217	10	US-09-815-242-12523	Sequence 12523, A
33	6	7.6	1531	10	US-09-876-889-347	Sequence 347, App
34	6	7.6	1531	10	US-09-998-598-2593	Sequence 2593, Ap
35	6	7.6	1569	9	US-10-108-605-303	Sequence 303, App
36	6	7.6	1601	10	US-09-862-027-40	Sequence 40, Appl
37	5	6.3	9	10	US-09-894-018-260	Sequence 260, App
38	5	6.3	15	10	US-09-879-957-150	Sequence 150, App
39	5	6.3	23	10	US-09-864-761-41290	Sequence 41290, A
40	5	6.3	24	10	US-09-864-761-40862	Sequence 40862, A
41	5	6.3	26	10	US-09-287-849-39	Sequence 39, Appl1
42	5	6.3	26	10	US-09-864-761-41246	Sequence 41246, A
43	5	6.3	27	10	US-09-864-761-43417	Sequence 43417, A
44	5	6.3	28	10	US-09-864-761-39263	Sequence 39263, A
45	5	6.3	30	10	US-09-864-761-36827	Sequence 36827, A

ALIGNMENTS

RESULT 1
US-09-864-761-44740
Sequence 44740, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 44740
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC016057.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52
; OTHER INFORMATION: SWISSPROT HIT: Q13563, EVALUATION = 2.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF570694.1, EVALUATION = 4.00e-16
US-09-864-761-44740
```

```

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 23 AHPVLSF 29
Db 37 AHPVLSF 43
```

```

RESULT 2
US-10-001-876-210
; Sequence 210, Application US/10001876
; Patent No. US20020177140A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Cafferey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0285
; CURRENT APPLICATION NUMBER: US/10/001,876
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,186
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 210
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-001-876-210
```

```

Query Match
Best Local Similarity 100.0%; Score 7; DB 9; Length 215;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 27 LSFRLPS 33
Db 60 LSFRLPS 66
```

```

RESULT 3
US-09-864-761-41953
; Sequence 41953, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 41953
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC023344.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
US-09-864-761-41953
```

```

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 47 PLSPTS 52
Db 24 PLSPTS 29
```

```
RESULT 4
```

US-09-764-887-226
; Sequence 226, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA113
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 226
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-887-226

Query Match 7.6%; Score 6; DB 10; Length 68;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 ISPVSV 60
|||||
Db 28 ISPVSV 33

RESULT 5
US-09-764-869-1118
; Sequence 1118, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1118
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-1118

Query Match 7.6%; Score 6; DB 10; Length 86;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 LPLAPL 48
|||||
Db 61 LPLAPL 66

RESULT 6
US-09-925-301-1470
; Sequence 1470, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1470
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (117)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (141)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (146)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (148)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (152)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (153)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (158)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1470

Query Match 7.6%; Score 6; DB 10; Length 168;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 PVLSPFL 30
|||||
Db 121 PVLSPFL 126

RESULT 7
US-10-024-579-16
; Sequence 16, Application US/10024579
; Patent No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Fridgde, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 16
;; LENGTH: 238
;; TYPE: PRT
;; ORGANISM: homo sapiens
US-10-024-579-16

Query Match
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 VLSFLR 31
|111111|
Db 87 VLSFLR 92

RESULT 8
US-10-024-579-14
; Sequence 14, Application US/10024579
; Patent No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1 Human Ion Channel-Related Proteins
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 257
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-024-579-14

Query Match
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 VLSFLR 31
|111111|
Db 87 VLSFLR 92

RESULT 9
US-10-024-579-12
; Sequence 12, Application US/10024579
; Patent No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1 Human Ion Channel-Related Proteins
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 264
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-024-579-12

Query Match 7.6%; Score 6; DB 12; Length 264;

Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 VLSFLR 31
|111111|
Db 113 VLSFLR 118

RESULT 10
US-09-808-602-23
; Sequence 23, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-23

Query Match
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 LPLAPL 48
|111111|
Db 20 LPLAPL 25

RESULT 11
US-09-808-602-25
; Sequence 25, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-808-602-25

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 267;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 43 LPLAPL 48
Db 20 LPLAPL 25

RESULT 12

US-10-024-579-10
; Sequence 10, Application US/10024579
; Patent No. US20020119522A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Gerhardt, Brenda
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020119522A1 Human Ion Channel-Related Proteins
; FILE REFERENCE: LEX-0274-USA
; CURRENT APPLICATION NUMBER: US/10/024,579
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: US 60/258,595
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 283
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-024-579-10

Query Match
Best Local Similarity 100.0%; Score 6; DB 12; Length 283;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 26 VLSEFLR 31
Db 113 VLSEFLR 118

RESULT 13
US-10-041-006A-7
; Sequence 7, Application US/10041006A
; Patent No. US20020168754A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; TITLE OF INVENTION: protease T
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/041,006A
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 290
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-041-006A-7

Query Match
Best Local Similarity 100.0%; Score 6; DB 9; Length 290;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 RPAAPV 13
Db 3 RPAAPV 8

RESULT 14
US-10-028-072-222
; Sequence 222, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gettitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang

; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28

PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627

PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 7 6%: Score 6; DB 9; Length 290;
Best local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RPAVP 13
Db 3 RPAVP 8

RESULT 15
US-10-040-655-7
; Sequence 7, Application US/10040655
; Patent No. US20020146805A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia

APPLICANT: Darrow, Andrew
APPLICANT: Qi, Jian-shen
TITLE OF INVENTION: DNA encoding the novel human serine
TITLE OF INVENTION: Protease T
FILE REFERENCE: ORT-1032
CURRENT APPLICATION NUMBER: US/10/040,655
CURRENT FILING DATE: 2002-01-07
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 290
TYPE: PRT
ORGANISM: Homo sapiens
US-10-040-655-7

Query Match 7.6%; Score 6; DB 12; Length 290;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8 RPAVP 13
|||
|||
Db 3 RPAVP 8

Search completed: January 14, 2003, 17:21:49
Job time : 4.69679 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:10:35 ; Search time 2.69679 Seconds

(without alignments)
1215.012 Million cell updates/sec

Title: US-09-506-079e-1

Perfect score: 79
Sequence: 1 GKHSHXPRPAAPVPPKXQP.....VGRGXDPDAHVAVLSRYEG 79

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	7	8.9	114 1 GON2_TUPGB	G05336 tupala glis
2	7	8.9	380 1 YMH7_CABEL	P34474 caenorhabdi
3	7	8.9	574 1 ZEST_DROME	P09356 dirosophila
4	7	8.9	635 1 SL56_HUMAN	Q92899 homo sapien
5	7	8.9	677 1 GCP3_MOUSE	P58854 mus musculu
6	7	8.9	907 1 GCP3_HUMAN	Q96CWS homo sapien
7	7	7.6	56 1 PSBK_PINTH	P41598 pinus thund
8	7	7.6	126 1 MF14_MAIZE	Q01900 zea mays (m
9	7	7.6	145 1 VP3_BPCHP	P19194 bacterioph
10	7	7.6	195 1 GSPJ_ECOLI	Q94576 escherichia
11	7	7.6	198 1 LMBV_CHICK	Q01636 gallus gall
12	7	7.6	204 1 LAF4_MOUSE	P51827 mus musculu
13	7	7.6	206 1 COX3_MOUSE	Q04442 bacillus fi
14	7	7.6	207 1 YPO_P_BACSU	P34183 bacillus su
15	7	7.6	209 1 ADPP_ECOLI	P36651 escherichia
16	7	7.6	215 1 CYB6_ODOSI	P49488 odontella s
17	7	7.6	220 1 ACPD_STRCO	Q951U6 streptomyce
18	7	7.6	220 1 BIOD_AOUAE	Q66832 aquilex aeo
19	7	7.6	239 1 PRRA_BOVIN	P18917 bos taurus
20	7	7.6	241 1 LAT_RAT	O70601 ratu mus norv
21	7	7.6	242 1 LAT_MOUSE	O54957 mus musculu
22	7	7.6	248 1 PT16_LYCES	Q04682 lycopersico
23	7	7.6	260 1 YH20_PASMU	Q96KAS pasteurilla
24	7	7.6	263 1 COO4_CABEL	P31428 caenorhabdi
25	7	7.6	269 1 NIEP_AZOCB	P23145 azotobacter
26	7	7.6	280 1 ARI1_MESAU	P50292 mesocricetu
27	7	7.6	290 1 MPN_HUMAN	Q94943 homo sapien
28	7	7.6	301 1 YF34_SYNY3	P74220 synechocyst
29	7	7.6	313 1 FIXB_ECOLI	P31574 escherichia
30	7	7.6	313 1 Y135_TREPA	O83171 treponema p
31	7	7.6	322 1 CYSL_HOMAM	P33277 homarus ame
32	7	7.6	344 1 LPXK_NEITMA	Q91V64 neisseria m
33	7	7.6	344 1 LPXK_NEITMB	Q9K047 neisseria m

34	6	7.6	348 1	YT35_STRPR	P20186 streptomyce
35	6	7.6	349 1	PTE1_YEAST	P41903 saccharomyc
36	6	7.6	352 1	RECA_CLOPE	P46666 clostridium
37	6	7.6	378 1	AROB_RHIO	Q98FV1 rhizobium 1
38	6	7.6	382 1	CRTY_PANAN	P21687 pantoea ana
39	6	7.6	383 1	CBG_FABIT	P23775 oryctolagus
40	6	7.6	386 1	CRTY_ERWHE	Q01331 erwina her
41	6	7.6	390 1	CHL_MLVGN	P23092 cas-n5-1 mu
42	6	7.6	392 1	TRB1_MERTH	Q27696 methanobact
43	6	7.6	398 1	YAIR_ECOLI	Q47536 escherichia
44	6	7.6	400 1	TRAF6_BACST	Q45618 bacillus st
45	6	7.6	402 1	DPP_SYNY3	P73881 synechocyst

ALIGNMENTS

RESULT 1	ID	Accession	Standard	PRF	Length
GON2_TUPGB	GON2_TUPGB	STANDARD	PRF	114 AA.	
AC	Q95336				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	15-DEC-1998 (Rel. 39, Last annotation update)				
DT	30-MAY-2000 (Rel. 39, Last annotation update)				
DE	Progonadoliberin II precursor [Contains: Gonadoliberin II (LHRH II) (Luteinizing hormone releasing hormone II) (Gonadotropin releasing hormone II) (GNRH II) (Luliberin II); GNRH-associated peptide II].				
GN	GNRH2				
OS	Tupala glis belongs to (Common tree shrew).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Scandentia; Tupaiidae; Tupala.				
OX	NCBI_TaxID=9396;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Hypothalamus;				
RA	MEDLINE=97079639; PubMed=8921350;				
RA	Kasten T.L., White S.A., Norton T.T., Bond C.T., Adelman J.P.,				
RA	Fernald R.D.;				
RT	"Characterization of two new preproGNRH mRNAs in the tree shrew:				
RT	first direct evidence for mesencephalic GNRH gene expression in a				
RT	placental mammal."				
RL	Gen. Comp. Endocrinol. 104:7-19(1996).				
CC	-1- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS; IT STIMULATES				
CC	THE SECRETION OF BOTH LUTEINIZING AND FOLLICLE-STIMULATING				
CC	HORMONES.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- TISSUE SPECIFICITY: MIDBRAIN.				
CC	-1- SIMILARITY: BELONGS TO THE GNRH FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: U63327; AB16838.1; -				
DR	InterPro: IPR002012; GNRH.				
DR	Pfam: PF00446; GNRH.1.				
DR	PROSITE: PS00473; GNRH.1.				
KW	cleavage on pair of basic residues; Hormone; Amidation; Signal.				
FT	SIGNAL	1	25		
FT	CHARIN	26	114		
FT	PEPTIDE	26	35		
FT	MOD_RES	39	114		
FT	MOD_RES	26	26		
FT	MOD_RES	35	35		
FT	SEQUENCE	114 AA;	12123 MW;		
SO	Best Local Similarity	8.9%;	Score 7;		
		100.0%;	Pred. No. 2.8;		
			Length 114;		

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 PRPAVP 13
 Db 105 PRPAVP 111

RESULT 2

YMH7_CAEEL STANDARD; PRT; 380 AA.
 ID YMH7_CAEEL
 AC P34474;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical 31.8 kDa protein F58A4.7 in chromosome III.
 GN F58A4.7
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RC MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smailson N., Smith A., Smith M., Sonhammer E., Staden R.,
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Watson R., Watson A., Weinstock L., Wilkinson-Spratt J.,
 RA Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RT Nature 368:32-38(1994).

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. STRONGEST, TO TRANSCRIPTION FACTOR AP-4.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: Z22179; CAAB0167.1; -;
 CC PIR: S40979; S40979.
 CC Wormpep: F58A4.7; CE01018.
 CC InterPro: IPR001092; HLH_basic.
 CC Pfam: PF00010; HLH_1.
 CC SMART: SM00353; HLH_1.
 CC PROSITE: PS00038; HLH_1; 1.
 CC PROSITE: PS00888; HLH_2; 1.
 CC Hypothetical protein; Transcription regulation; DNA-binding;
 CC Nuclear protein.
 FT DNA_BIND 61 73 BASIC DOMAIN
 FT DOMAIN 74 113 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 SQ SEQUENCE 380 AA; 40519 MW; EE2B52DIEB3EPD33 CRC64;

Query Match 8.9%; Score 7; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 APLSPS 52
 Db 48 APLSPS 54

RESULT 3

ZEST_DROME STANDARD; PRT; 574 AA.
 ID ZEST_DROME
 AC P09956; O24596; Q9V3F1;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Regulatory protein zeste.
 GN Z OR EG:BACH59J11.3 OR CG7803.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88142560; PubMed=3125410;
 RA Mansukhani A., Gunaratne P.H., Sherwood P.W., Sneath B.J.,
 RA Goldberg M.L.;
 RT "Nucleotide sequence and structural analysis of the zeste locus of
 RT Drosophila melanogaster.";
 RT Mol. Gen. Genet. 211:121-128(1988).

RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon R;
 RX MEDLINE=87218538; PubMed=3582372;
 RA Piroletta V., Manet E., Hardon E., Bickel S.E., Benson M.;
 RT "Structure and sequence of the Drosophila zeste gene.";
 RT EMBO J. 6:791-799(1987).

RN (3)
 RP REVISIONS, AND SELF-ASSOCIATION.
 RX MEDLINE=90361011; PubMed=2118108;
 RA Bickel S.E., Piroletta V.;
 RT "Self-association of the Drosophila zeste protein is responsible for
 RT transvection effects.";
 RT EMBO J. 9:2959-2967(1990).

RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrel L., Fierla S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei L., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Pui Y., Reese M.G.,
 RA Palzer K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Oregon-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,
 RA Dreano S., Gloux S., Delaure V., Motlier S., Galibert F., Borkova D.,
 RA Minano B., Kafatos F.C., Louis C., Siden-Klamos I., Bolshakov S.,
 RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
 RA Modelall J., Peter A., Schoettler P., Werner M., Mourikioti F.,
 RA Belnert N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamsou A., Henderson N.S.,
 RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;
 RT "From sequence to chromosome: the tip of the X chromosome of *D.*
 RT *melanogaster*.";
 RL Science 287:2220-2232(2000).
 RN [6]
 RP SELF-ASSOCIATION.
 RX MEDLINE=93259149; PubMed=8491197;
 RA Chen J.D., Piroletta V.;
 RT "Multimerization of the *Drosophila* zeste protein is required for
 RT efficient DNA binding.";
 RL EMBO J. 12:2075-2083(1993).
 RN [7]
 RP SEQUENCE OF 56-323 FROM N.A.
 RC STRAIN-ME-K1, ME-K2, ME-L1L, ME-L1Z, ME-NJ1, and ME-NJ2;
 RX MEDLINE=93360802; PubMed=8355601;
 RA Hey J., Kilman R.M.;
 RT "Population genetics and phylogenetics of DNA sequence variation at
 RT multiple loci within the *Drosophila melanogaster* species complex.";
 RL Mol. Biol. Evol. 10:804-822(1993).
 CC -1- FUNCTION: INVOLVED IN TRANSECTON PHENOMENA ("SYNOPSIS-DEPENDENT
 CC CARRYING GENES WITH WHICH ZESTE INTERACTS INFLUENCES THE
 CC EXPRESSION OF THESE GENES. ZESTE BINDS TO DNA AND STIMULATES
 CC TRANSCRIPTION FROM A NEARBY PROMOTER.
 CC -1- SUBUNIT: SELF-ASSOCIATE FORMING COMPLEXES OF SEVERAL HUNDRED
 CC MONOMERS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Y00049; CAA68262.1; ALT_SEQ.
 DR EMBL: X06743; CAA29918.1; ALT_SEQ.
 DR EMBL: L13043; AAA29026.1; -;
 DR EMBL: L13044; AAA29027.1; -;
 DR EMBL: L13045; AAA29028.1; -;
 DR EMBL: L13046; AAA29029.1; -;
 DR EMBL: L13047; AAA29030.1; -;
 DR EMBL: L13048; AAA29031.1; -;
 DR EMBL: AE003424; AAF45783.1; -;
 DR EMBL: AL133505; CAB63525.1; -;
 DR PIR: A26639; A26639.
 DR PIR: S01272; S01272.
 DR PIR: S12369; S12369.
 DR TRANSFAC: T00918; -;
 DR FlyBase: FBgn0004050; z.
 KW DNA-binding; Transcription regulation; Nuclear protein; Polymorphism.
 FT INT_MET 0
 FT DONAIN 1 46 HYDROPHOBIC.
 FT DNA_BIND 47 127 SPECIFIC, WITH ZESTE LOCUS.

FT DOMAIN 5 42 GLY-RICH.
 FT DOMAIN 152 430 GLN/ALA-RICH (OPA-REPEAT INVOLVED IN
 FT TRANSCRIPTIONAL ACTIVATION OR REPRESSION
 FT AT DIFFERENT TARGET LOCUS) (POTENTIAL).
 FT VARIANT 232 233 MISSING (IN STRAIN ME-K2).
 FT CONFLICT 496 496 S -> A (IN REF. 4 AND 5).
 SQ SEQUENCE 574 AA; 61845 MM; 45DCD36CE72F5CF3 CRC64;
 Query Match 8.9%; Score 7; DB 1; Length 574;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 PAAVPVP 15
 DB 449 PAAVPVP 455
 RESULT 4
 SL56_HUMAN STANDARD: PRT; 635 AA.
 ID SL56_HUMAN
 AC Q9Y289;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sodium-dependent multivitamin transporter (Na⁺)-dependent
 DE multivitamin transporter).
 GN SLC5A6 OR SMVT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Intestine;
 RX MEDLINE=99262640; PubMed=10329687;
 RA Wang H., Huang W., Fel Y.-J., Xia H., Yang-Feng T.L., Leibach F.H.,
 RA Devoe L.D., Ganapathy V., Prasad P.D.;
 RT "Human placental Na⁺-dependent multivitamin transporter. Cloning,
 RT functional expression, gene structure, and chromosomal localization.";
 RL J. Biol. Chem. 274:14875-14883(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Intestine;
 RX MEDLINE=99268779; PubMed=10334869;
 RA Prasad P.D., Wang H., Huang W., Fel Y.-J., Leibach F.H., Devoe L.D.,
 RA Ganapathy V.;
 RT "Molecular and functional characterization of the intestinal Na⁺-
 RT dependent multivitamin transporter.";
 RL Arch. Biochem. Biophys. 366:95-106(1999).
 CC -1- FUNCTION: TRANSPORTS PANTOTHENATE, BIOTIN AND LIPDATE IN THE
 CC PRESENCE OF SODIUM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:SOLUTE SYMPORTER FAMILY (SSF).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF116241; AAD37502.1; -;
 DR EMBL: AF069307; AAD31727.1; -;
 DR EMBL: AF081571; AAD37481.1; -;
 DR Genew; HGNC:11041; SLC5A6.
 DR MIM: 604024; -;
 DR InterPro: IPR001734; Na/solut_symport.
 DR Pfam: PF00474; SSF; 1.
 DR TIGRFAMs: TIGR00813; sss; 1.
 DR PROSITE: PS00456; NA_SOLUT_SYMP_1; 1.
 DR PROSITE: PS00457; NA_SOLUT_SYMP_2; FALSE_NEG.
 DR PROSITE: PS50283; NA_SOLUT_SYMP_3; 1.

KW Transport; Transmembrane; Sodium transport; Symport; Glycoprotein.
 FT TRANSMEM 24 44 POTENTIAL.
 FT TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 101 121 POTENTIAL.
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 176 196 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 336 356 POTENTIAL.
 FT TRANSMEM 396 416 POTENTIAL.
 FT TRANSMEM 428 448 POTENTIAL.
 FT TRANSMEM 456 476 POTENTIAL.
 FT TRANSMEM 528 548 POTENTIAL.
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 635 AA; 68701 MW; 4F10369A3916564F CRC64;

Query Match 8.9%; Score 7; DB 1; Length 635;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 APUSPTS 52
 DB 9 APUSPTS 15

RESULT 5
 GCP3_MOUSE STANDARD; PRT; 677 AA.
 AC P58854;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Gamma-tubulin complex component 3 (GCP-3) (Fragment).
 GN GCP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule
 nucleation at the centrosome (By similarity).
 CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,
 GCP3, GCP4, GCP5 and GCP6 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Centrosome (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: BC025647; AAH25647.1; -.
 KW Microtubules.
 FT NON_TER 1 1
 FT DOMAIN 595 599 POLY-GIU.
 SQ SEQUENCE 677 AA; 78348 MW; 16A85A47357E33C4 CRC64;

Query Match 8.9%; Score 7; DB 1; Length 677;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 HPVUSFL 30
 DB 11111111

DB 194 HPVUSFL 200

RESULT 6
 GCP3_HUMAN STANDARD; PRT; 907 AA.
 AC G96CW5; O43631; O60852; O60853; O96179;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Gamma-tubulin complex component 3 (GCP-3) (Spindle pole body protein
 Spc98 homolog) (hGCP3) (h104p).
 GN GCP3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE=98234402; PubMed=9566967;
 RA Murphy S.M., Urban L., Stearns T.;
 RT "The mammalian gamma-tubulin complex contains homologues of the yeast
 spindle pole body components spc97p and spc98p.";
 RL J. Cell Biol. 141:663-674(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
 RX MEDLINE=98234404; PubMed=9566969;
 RA Tassin A.-M., Celati C., Moudjou M., Bornens M.;
 RT "Characterization of the human homologue of the yeast spc98p and its
 association with gamma-tubulin.";
 RL J. Cell Biol. 141:689-701(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
 RC TISSUE=Eye, and Muscle;
 RA Strausberg R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule
 nucleation at the centrosome.
 CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,
 GCP3, GCP4, GCP5 and GCP6.
 CC -1- SUBCELLULAR LOCATION: Centrosome.
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3, may be
 produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF042378; AAC39727.1; -.
 DR EMBL: AJ003061; CAA05832.1; -.
 DR EMBL: AJ003062; CAA05833.1; -.
 DR EMBL: BC013781; AAH13781.1; -.
 DR EMBL: BC007763; AAH07763.1; -.
 KW Microtubules; Alternative splicing.
 FT DOMAIN 825 829
 FT VARSPLIC 817 824 GOMGVNAA -> VEMCLYCV (IN ISOFORM 2).
 FT VARSPLIC 825 907 MISSING (IN ISOFORM 2).
 FT VARSPLIC 391 434 RKGELASVHATKRGDPMRSTNOHIIISVSHPTLSLY
 RMI -> PIRVPTFHVPTPRDPTPTDPMHVPTRVPTRV
 WHSLCFRTLL (IN ISOFORM 3).
 FT VARSPLIC 435 907 MISSING (IN ISOFORM 3).
 FT CONFLICT 208 208 T -> S (IN REF. 2).
 FT CONFLICT 361 361 S -> I (IN REF. 2).
 SQ SEQUENCE 907 AA; 103570 MW; 70FE2FDB7680344D CRC64;

Query Match 8.9%; Score 7; DB 1; Length 907;

Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 HPVLSFL 30
|||||||
Db 424 HPVLSFL 430

RESULT 7

PSBK_P1NTH STANDARD; PRT: 56 AA.

ID PSBK_P1NTH
AC P41598;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II reaction center protein K precursor (PSII-K).
GN PSBK.
OS Pinus thunbergii (Green pine) (Japanese black pine).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
NCBI_TaxID=3350;

[1]
SEQUENCE FROM N.A.
RX MEDLINE=95024047; PubMed=7937893;
RA Wakasugi T., Tsudzuki J., Ito S., Nakashima K., Tsudzuki T.,
Sugitara M.;
RT "Loss of all ndh genes as determined by sequencing the entire
chloroplast genome of the black pine Pinus thunbergii.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9794-9798(1994).
CC -1- FUNCTION: THIS PROTEIN IS A COMPONENT OF THE REACTION CENTER
OF PHOTOSYSTEM II.
CC -1- SIMILARITY: BELONGS TO THE PSBK FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; D17510; BAA04312.1; ALT_INIT.
DR InterPro; IPR003687; PSII_PSDK.
DR Pfam; PF02533; Psbk; 1.
KW Photosystem II; Chloroplast.
FT PROPEP 1 19 POTENTIAL.
FT CHAIN 20 56 PHOTOSYSTEM II REACTION CENTER PROTEIN K.
SQ SEQUENCE 56 AA; 6346 MW; 18BIDEF198ACA5A CRC64;

Query Match 7.6%; Score 6; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 PVLISFL 30
|||||||
Db 39 PVLISFL 44

RESULT 8
MFL4_MAIZE STANDARD; PRT: 126 AA.

ID MFL4_MAIZE
AC O01900;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE MFL4 protein precursor.
GN MFL4.

OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Palcoideae; Andropogoneae; Zea.
NCBI_TaxID=4577;

[1]
SEQUENCE FROM N.A.

RP STRAIN=CV. BEL0; TISSUE=Tassel;
RX MEDLINE=94004987; PubMed=8401606;

RA Wright S.Y., Sener M.-M., Bell P.J., Vaudin M., Greenland A.J.;
RT "Isolation and characterization of male flower cDNAs from maize.";
RL Plant J. 3:41-49(1993).

CC -1- TISSUE SPECIFICITY: ENHANCED EXPRESSION IN MALE FLOWERS.
CC ACCUMULATES IN THE TAPETUM.
CC -1- DEVELOPMENTAL STAGE: ASSOCIATED WITH MICROSPOROGENESIS AND
CC DECLINES AS MATURE POLLEN IS PRODUCED.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X67323; CAA47737.1; -
DR PIR; S25104; S25104.

DR MaizeDB; 69183; -

KW Signal.
FT SIGNAL 1 23 OR 24, OR 26 (POTENTIAL).
FT CHAIN 24 126 MFL4 PROTEIN.
SQ SEQUENCE 126 AA; 12653 MW; 67F2813AF8C55E5 CRC64;

Query Match 7.6%; Score 6; DB 1; Length 126;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 RPAAPV 13
|||||||
Db 38 RPAAPV 43

RESULT 9
VP3_BPCHP STANDARD; PRT: 145 AA.

ID VP3_BPCHP
AC P19194;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE Protein VP3 (ORF3).
OS Bacteriophage Chpl.
OC Viruses; ssDNA viruses; Microviridae; Microvirus.
NCBI_TaxID=12367;

[1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90111716; PubMed=2607341;
RA Storey C.C., Lusher M., Richmond S.J.;
RT "Analysis of the complete nucleotide sequence of Chpl, a phage which
RT infects avian Chlamydia psittaci.";
RL J. Gen. Virol. 70:3381-3390(1989).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D00624; BAA00509.1; -
DR PIR; J00347; J00347.

KW structural protein.
SQ SEQUENCE 145 AA; 16681 MW; 6B4DFA6463587CD CRC64;

Query Match 7.6%; Score 6; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 39 AFYSLP 44
DB 83 AFYSLP 88

RESULT 10
GSPJ_ECOLI
ID GSPJ_ECOLI STANDARD; PRT; 195 AA.
AC P45761;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable general secretion pathway protein J precursor.
GN GSPJ OR B3331.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
CC EXPORT OF PROTEINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PULJ/OUTJ/XRPSJ/EXEL/ACPM FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U18997; AAA58128.1; -
DR EMBL: AE000409; AAC76356.1; -
DR Ecogene: EG12893; gspJ.
DR InterPro: IPR001120; Prok_N.methyln.
DR PROSITE: PS00409; PROKAR_NTER_METHYL; 1.
KW Transport; Methylation; Complete proteome.
FT PROPEP 1 7
FT CHAIN 8 195
FT MOD_RES 8 8 METHYLATION (BY SIMILARITY).
FT SEQUENCE 195 AA; 22231 MW; 4A25DB082106B1C5 CRC64;
SQ
Query Match 7.6%; Score 6; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 PLAPLS 49
DB 92 PLAPLS 97

RESULT 11
LMHV_CHICK
ID LMHV_CHICK STANDARD; PRT; 198 AA.
AC O01636;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin beta-1 chain variant (Laminin beta-1-2 chain) (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

```

```

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Eye;
RC MEDLINE=93015947; PubMed=1400373;
RA O'Rear J.J.;
RT "A novel laminin B1 chain variant in avian eye.";
RL J. Biol. Chem. 267:20555-20557(1992).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration, and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L00963; AAA9140.1; -
DR InterPro: IPR01886; LamNT.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF00055; laminin_Nterm; 1.
DR ProDom: PD002082; LamNT; 1.
DR SMART: SM00136; LamNT; 1.
DR PROSITE: PS01248; LAMININ_TYPE_EGF_PARTIAL.
KW Extracellular matrix; Cell adhesion; Glycoprotein; Basement membrane;
KW Laminin EGF-like domain.
FT NON_TER 1 1
FT DOMAIN 74 >198 LAMININ N-TERMINAL (DOMAIN VI).
FT NON_TER 198 198
SQ SEQUENCE 198 AA; 21830 MW; 6FD669761892C442 CRC64;
SQ
Query Match 7.6%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PRAAV 12
DB 59 PRAAV 64

RESULT 12
LAF4_MOUSE
ID LAF4_MOUSE STANDARD; PRT; 204 AA.
AC P51827;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE LAF-4 protein (Lymphoid nuclear protein related to AFA) (Fragment).
GN LAF4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96141096; PubMed=8555498;
RA Ma C., Staedt L.M.;
RT "LAF-4 encodes a lymphoid nuclear protein with transactivation
RT potential that is homologous to AFA-4, the gene fused to MLL in
RT t(4;11) leukemias.";

```



```

RL Blood 87:734-745(1996).
CC -1- FUNCTION: POTATIVE TRANSCRIPTION ACTIVATOR THAT MAY FUNCTION IN
CC LYMPHOID DEVELOPMENT AND ONCOGENESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN LYMPHOID TISSUES,
CC LOWER LEVELS IN BRAIN AND LUNG.
CC -1- SIMILARITY: BELONGS TO THE AP4 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U34361; AAA98764.1; -.
DR MED: MGI:106927; Laf4.
DR NUCLEIC ACID: Transcription regulation; Activator; DNA-binding.
DR NON_TER 204
DR SEQUENCE 204 AA; 22434 MW; CA1FC61D46477F59 CRC64;

Query Match
Best Local Similarity 7.6%; Score 6; DB 1; Length 204;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PAAVPV 14
Db 133 PAAVPV 138

RESULT 13
COX3_BACFI STANDARD; PRT; 206 AA.
ID COX3_BACFI
AC 004442;
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytochrome c oxidase polypeptide III (EC 1.9.3.1) (Cytochrome Aa3
DE subunit 3).
DE CTRAE.
OS Bacillus firmus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID:1399;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OF4;
RX MEDLINE-93107080; PubMed-7678007;
RA Quirk P.G., Hicks D.B., Krulwich T.A.;
RA "Cloning of the cta operon from alkaliphilic Bacillus firmus OF4 and
RA characterization of the pH-regulated cytochrome caa3 oxidase it
RA encodes."
RL J. Biol. Chem. 268:678-685(1993).
CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) -> 4 ferricytochrome
CC c + 2 H(2)O.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M94110; AAA22366.1; -.
DR PIR: E45335; E45335.
DR InterPro: IPR000298; CytC_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PRODOM: PD000382; CytC_oxdse_III; 1.
DR PROSITE: PS50253; COX3; 1.
DR Oxidoreductase; Transmembrane.

```

```

FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 97 117 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
SO SEQUENCE 206 AA; 22856 MW; B30CA3D339FC4483 CRC64;

Query Match
Best Local Similarity 7.6%; Score 6; DB 1; Length 206;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 SAFYSL 43
Db 138 SAFYSL 143

RESULT 14
YPOP_BACSU STANDARD; PRT; 207 AA.
ID YPOP_BACSU
AC P54183;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ypop (fragment).
DE YPOP.
GN Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID:1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168 / Marburg;
RX MEDLINE-96349105; PubMed-8760912;
RA Sorokin A.V., Azevedo V., Zumbstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RA "Sequence analysis of the Bacillus subtilis chromosome region between
RA the sera and kdg loci cloned in a yeast artificial chromosome."
RL Microbiology 142:2005-2016(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE-98044033; PubMed-9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertsch M.G., Bessieres P., Bolyard A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brigelli S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabre C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppl G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerli-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kunita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mosti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porroli S., Prescott A.M.,
RA Priesen E., Puig C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenol M., Vannier F., Vassarelli A.,
RA Viari A., Wambuit R., Wedler E., Wedler H., Wellzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus
RA subtilis."
RL Nature 390:249-256(1997).
CC -1- CAUTION: This ORF is interrupted by insertion of the SPbeta

```

```

CC -----
CC      prophage.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL: L77246; AAA96649.1; -
CC      EMBL: Z99115; CAB14085.1; -
CC      Subtilist; BG11629; yppp
CC      InterPro: IPR003869; Polysac_CapD
CC      Pfam: PF02719; Polysacc_synth_2; 1.
CC      Hypothetical protein; Complete proteome.
CC      NON_TER
CC      SEQUENCE 207 AA; 23286 MW; F3EA64945DE249CD CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 207;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PRPAV 12
Db 126 PRPAV 131

RESULT 15
ADPP_ECOLI STANDARD; PRT; 209 AA.
ID ADPP_ECOLI
AC P36651;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ADP-ribose pyrophosphatase (EC 3.6.1.13) (ADP-ribose diphosphatase)
DE (Adenosine diphosphoribose pyrophosphatase) (ADPR-Prase) (ADP-ribose
DE phosphohydrolase).
GN NUDE OR B3034 OR Z4391 OR ECS3922.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OC NCBI_TaxID=562, 83334;
OX
RN
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=1120551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamodis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN
RN SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,

```

```

RA Kihara S., Shida T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN
RN SEQUENCE OF 98-209 FROM N.A.
RC STRAIN=K12;
RA Imamura R., Niki H., Yamanaka K., Ogura T., Fujita N., Ishihama A.,
RA Hiraga S.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
RN
RN CHARACTERIZATION.
RC STRAIN=K12 / K61655;
RX MEDLINE=20011437; PubMed=10542272;
RA Dunn C.A., O'Handley S.F., Frick D.N., Bessman M.J.;
RT "Studies on the ADP-ribose pyrophosphatase subfamily of the nudix
RT hydrolases and tentative identification of tngb, a gene associated
RT with tellurite resistance.";
RL J. Biol. Chem. 274:32318-32324(1999).
CC
CC -I- CATALYTIC ACTIVITY: ADP-ribose + H(2)O = AMP + D-ribose 5-
CC phosphate.
CC -I- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY. NUDE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U28377; AAA69202.1; -
CC DR AEMBL: AE000385; AAC76070.1; -
CC DR EMBL: AE005533; AAG58173.1; -
CC DR EMBL: AP002564; BAB37345.1; -
CC DR EMBL: D16557; -; NOT_ANNOTATED_CDS.
CC DR Ecogene: EG12184; nude.
CC DR InterPro: IPR004385; Cons_hypoth52.
CC DR Pfam: PF00293; NUDIX; 1.
CC DR PRINTS: PR00502; NUDIXFAMILY.
CC DR TIGRfam: TIGR00052; Cons_hypoth52; 1.
CC DR PROSITE: PS00893; NUDIX; 1.
CC KW Hydrolase; Magnesium; Manganese; Zinc; Complete proteome.
CC FT DOMAIN 97 118 NUDIX BOX.
CC SEQUENCE 209 AA; 23667 MW; 2CF7EA9D63B9615 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 209;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PVLSFL 30
Db 126 PVLSFL 131

```

Search completed: January 14, 2003, 17:13:20
 Job time : 5.69679 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:11:37 ; Search time 4.6004 Seconds

(without alignments)
1650,860 Million cell updates/sec

Title: US-09-506-079E-1

Perfect score: 79

Sequence: 1 GXHSXPAPPAVPVPRXRP.....VGRGXDPDAHYVXLSRYEG 79

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR_73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Seq#	Query Match	Length	DB ID	Description
1	7	8.9	121	2 AH1848	hypothetical prote
2	7	8.9	292	2 S40979	hypothetical prote
3	7	8.9	382	2 B88561	protein F58A.7b l
4	7	8.9	389	2 A97577	hypothetical prote
5	7	8.9	389	2 AH2797	conserved hypotet
6	7	8.9	436	2 B70473	protoporphyrinogen
7	7	8.9	495	2 T17478	hypothetical prote
8	7	8.9	520	2 C70311	hypothetical prote
9	7	8.9	555	2 A26639	regulatory protein
10	7	8.9	574	2 S01272	regulatory protein
11	7	8.9	595	2 T15862	hypothetical prote
12	7	8.9	865	2 A85032	hypothetical prote
13	7	8.9	879	2 S49910	hypothetical prote
14	7	8.9	1503	2 T01098	chloroplast outer
15	7	8.9	2240	2 T37057	probable multi-dom
16	7	7.6	59	2 T07432	photosystem II pro
17	7	7.6	65	2 C72538	hypothetical prote
18	7	7.6	74	2 F82764	hypothetical prote
19	7	7.6	98	2 S68136	NADH2 dehydrogenas
20	7	7.6	120	2 D83166	hypothetical prote
21	7	7.6	122	2 AB2199	hypothetical prote
22	7	7.6	126	2 S25104	hypothetical prote
23	7	7.6	128	2 G71264	conserved hypotet
24	7	7.6	128	2 A87510	hypothetical prote
25	7	7.6	129	2 B83173	hypothetical prote
26	7	7.6	142	2 C58723	hypothetical prote
27	7	7.6	142	2 E72571	hypothetical prote
28	7	7.6	145	2 J00347	capsid protein VP3
29	7	7.6	163	2 AB2025	hypothetical prote

30	6	7.6	171	2 T43959	hypothetical prote
31	6	7.6	182	2 S76346	hypothetical prote
32	6	7.6	182	2 F69221	hypothetical prote
33	6	7.6	191	2 T04853	hypothetical prote
34	6	7.6	195	2 F65126	probable general s
35	6	7.6	198	2 A45067	lamnin BI chain v
36	6	7.6	205	2 F85515	unknown protein en
37	6	7.6	205	2 C90665	cytochrome-c oxida
38	6	7.6	206	2 E45335	capsular polysacch
39	6	7.6	207	2 A69941	ADPribiose diphosph
40	6	7.6	209	2 A85964	ADPribiose diphosph
41	6	7.6	209	2 B91119	ADPribiose diphosph
42	6	7.6	209	2 H65090	hypothetical prote
43	6	7.6	209	2 T21742	hypothetical prote
44	6	7.6	214	2 T29080	probable transcrip
45	6	7.6	215	2 S78307	plastoquinol-plast

ALIGNMENTS

RESULT 1
AH1848
hypothetical protein al10337 [Imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AH1848
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Itigun Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH1848
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-121 <KUR>
A:Cross-References: GB:BA000019; PIDN:BA872295.1; PID:q17129682; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: al10337

Query Match 8.9%; Score 7; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PVLSPFR 31
DB 56 PVLSPFR 62

RESULT 2

hypothetical protein F58A.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Feb-1997

A:Accession: S40979

R:Berts, M. submitted to the EMBL Data Library, February 1992

A:Reference number: S40973

A:Accession: S40979

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-292 <BER>

A:Cross-References: EMBL:Z22179

C:Genetics:

A:introns: 60/2; 133/1; 273/1

C:Keywords: nucleus

Query Match 8.9%; Score 7; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 APLSPTS 52
DB 48 APLSPTS 54

RESULT 3

B88561
protein F58A.7b [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B88561
R:anonymous, The C. elegans Sequencing Consortium.
A:title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MIMD:99069613; PMID:9651916
A:Note: see webstles genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C_elegans/; published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88561
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA80170.1; PID:g3877829; GSPDB:GN00021; CESP:F58A4
C:Genetics:
A:Gene: F58A.7b
A:Map position: 3

Query Match 8.9%; Score 7; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 APLSPTS 52
DB 48 APLSPTS 54

RESULT 4

A97577
hypothetical protein AGR_C_3311 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: A97577
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ounullo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: A97577
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-389 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK87570.1; PID:g15156910; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_3311
A:Map position: circular chromosome

Query Match 8.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVVP 15
DB 160 PAAVVP 166

RESULT 5

AH2797
conserved hypothetical protein Atu1800 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AH2797
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, T.
erge, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayvan, T.; Levy, R.; Li, M.; McCell
; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
ster, B.W.

A:title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AH2797
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-389 <KUR>
A:Cross-references: GB:AE008688; PIDN:AAL42798.1; PID:g17740243; GSPDB:GN00186
C:Experimental source: strain C58 (Dupont)
A:Genetics:
A:Gene: Atu1800
A:Map position: circular chromosome

Query Match 8.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAVVP 15
DB 160 PAAVVP 166

RESULT 6

B70473
protoporphyrinogen oxidase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: B70473
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MIMD:98196666; PMID:9537320
A:Accession: B70473
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-436 <AQF>
A:Cross-references: GB:AE000768; NID:g2984249; PIDN:AAC07778.1; PID:g2984251; GB:AE00
C:Genetics:
A:Gene: hemg

Query Match 8.9%; Score 7; DB 2; Length 436;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 VLSFLRP 32
DB 428 VLSFLRP 434

RESULT 7

T17478
hypothetical protein PCZA361.11 - Amycolatopsis orientalis
C:Species: Amycolatopsis orientalis
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T17478
R:Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard
Chem. Biol. 3, 155-162, 1998
A:title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin
A:Reference number: Z18804
A>Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-495 <VAN>
A:Cross-references: EMBL:AJ223998; NID:e1251208; PID:e1251217; PIDN:CAA11769.1

Query Match 8.9%; Score 7; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 PRPAVP 13
|||||||
Db 27 PRPAVP 33

RESULT 8

hypothetical protein aq_116 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jun-2000
C:Accession: C70311
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: C70311
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-520 <ADP>
A:Cross-references: GB:AE000674; NID:g2982850; PIDN:AC06482.1; PID:g2982861; GB:AE00065
A:Experimental source: strain VPS
C:Genetics:
A:Gene: aq_116
C:Superfamily: Aquifex aeolicus hypothetical protein aq_116

Query Match 8.9%; Score 7; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 PLSPRSV 53
|||||||
Db 2 PLSPRSV 8

RESULT 9

regulatory protein zeste - fruit fly (Drosophila sp.)
C:Species: Drosophila sp.
C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 16-Feb-1997
C:Accession: A26639
R:Pirota, V.; Manet, E.; Hardon, E.; Bickel, S.E.; Benson, M.
EMBO J. 6, 791-799, 1987
A:Title: Structure and sequence of the Drosophila zeste gene.
A:Reference number: A26639; MUID:87218538; PMID:3582372
A:Accession: A26639
A:Molecule type: DNA
A:Residues: 1-555 <PIR>
C:Genetics:
A:Gene: zeste
A:Cross-references: FlyBase:FBgn0004050
A:Introns: 103/2; 284/1
C:Keywords: DNA binding; transcription regulation

Query Match 8.9%; Score 7; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAPVP 15
|||||||
Db 430 PAAPVP 436

RESULT 10

regulatory protein zeste - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jul-2000
C:Accession: S01272
R:Mansuetti, A.; Gunaratne, P.H.; Sherwood, P.W.; Sneath, B.J.; Goldberg, M.L.
Mol. Gen. Genet. 211, 121-128, 1988
A:Title: Nucleotide sequence and structural analysis of the zeste locus of Drosophila me
A:Reference number: S01272; MUID:88142560; PMID:3125410

A:Accession: S01272
A:Molecule type: DNA
A:Residues: 1-574 <MAN>
A:Cross-references: EMBL:X06743; NID:g8852; PIDN:CAA29918.1; PID:g8853
A:Note: the authors translated the codon CGT for residue 366 as Ala
C:Genetics:
A:Gene: zeste
A:Cross-references: FlyBase:FBgn0004050
A:Introns: 123/2; 304/1
C:Keywords: DNA binding; transcription regulation

Query Match 8.9%; Score 7; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAPVP 15
|||||||
Db 449 PAAPVP 455

RESULT 11

hypothetical protein C56E6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15862
R:Fulton, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C56E6.
A:Reference number: S69019
A:Accession: T15862
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-595 <FUJ>
A:Cross-references: EMBL:U39996; NID:g1055114; PID:g1055118; PIDN:AAA81092.1; CESP:CS
A:Genetics:
A:Gene: CESP:C56E6.5
A:Introns: 23/2; 88/1; 134/3; 174/1; 265/1; 293/3; 362/2; 392/1; 436/2; 494/2; 539/2

Query Match 8.9%; Score 7; DB 2; Length 595;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 VLSFLRP 32
|||||||
Db 357 VLSFLRP 363

RESULT 12

hypothetical protein AT4G02510 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: A85032
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: A85032
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-865 <STO>
A:Cross-references: GB:NC_001268; NID:g7269011; PIDN:CAB80744.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4G02510
A:Map position: 4

Query Match 8.9%; Score 7; DB 2; Length 865;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 PAAPVP 15
|||||||

Db 584 PAAVPVP 590

RESULT 13

S49910

Chloroplast outer envelope protein OEP86 precursor - garden pea
N:Alternate names: chloroplast import-associated protein IAP86, GTP-binding
C/Species: Pisum sativum (garden pea)
C/Date: 26-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 02-Feb-2001
C/Accession: S49910; A55386; A55171
R:Hirsch, S.; Soll, J.
Submitted to the EMBL Data Library, March 1994
A/Reference number: S49910
A/Accession: S49910
A:Molecule type: mRNA
A:Residues: 1-879 <HIR>
A/Cross-references: EMBL:Z31581; NID:9599957; PID:9599958
R:Hirsch, S.; Muckel, E.; Heemeyer, F.; von Heijne, G.; Soll, J.
Science 266, 1989-1992, 1994
A/Title: A receptor component of the chloroplast protein translocation machinery.
A/Reference number: A55386; MID:95099324; PMID:7801125
A/Accession: A55386
A/Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-137 'Q', 139-875, 877-879 <HI2>
R:Kestler, F.; Blobel, G.; Patel, H.A.; Schnell, D.J.
Science 266, 1035-1039, 1994
A/Title: Identification of two GTP-binding proteins in the chloroplast protein import ma
A/Reference number: A55171; MID:95063938; PMID:7973656
A/Accession: A55171
A:Molecule type: mRNA
A:Residues: 1-372, 'EQQ', 376-879 <RES>
A/Cross-references: GB:I36857; NID:9576508; PID:AAA53276.1; PID:9576509
C/Keywords: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop
F:245-252/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 100.0%; Pred. No. 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PAAVPVP 15

Db 599 PAAVPVP 605

RESULT 14

T01098

Chloroplast outer envelope protein OEP86 homolog T10P11.19 - Arabidopsis thaliana
N:Alternate names: chloroplast import-associated protein, GTP-binding; protein T14P8.24
C/Species: Arabidopsis thaliana (mouse ear cress)
C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Feb-2001
C/Accession: T01098; T01299
R:Kaplan, N.; Johnson, D.; Schutz, K.; Guo, L.; Hoffman, J.; Till, S.; de la Bastide, M.
hl, M.; Martienssen, R.; Chen, E.Y.; Wilson, R.; McCombie, W.R.
Submitted to the EMBL Data Library, November 1998
A/Description: Sequence of A. thaliana BAC T10P11 from chromosome IV.
A/Reference number: Z14248
A/Accession: T01098
A/Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1503 <KAP>
A/Cross-references: EMBL:AC002330; NID:92262135; PID:93892053
A/Experimental source: cultivar Columbia
R:Kalkicki, J.; Elliott, G.; Cloud, J.
Submitted to the EMBL Data Library, May 1998
A/Description: The sequence of A. thaliana T14P8.
A/Reference number: Z14290
A/Accession: T01299
A/Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1503 <KAL>
A/Cross-references: EMBL:AF069298; NID:93193282; PID:93193301
A/Experimental source: cultivar Columbia

C/Genetics:
A:Map position: 4
A:Introns: 22/1

A/Note: T10P11.19; T14P.24
C/Keywords: chloroplast; GTP binding; membrane protein; nucleotide binding; P-loop
F:862-869/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 100.0%; Pred. No. 53;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PAAVPVP 15

Db 1222 PAAVPVP 1228

RESULT 15

T37057

Probable multi-domain beta keto-acyl synthase - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C/Accession: T37057
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
Submitted to the EMBL Data Library, August 1999
A/Reference number: Z21620
A/Accession: T37057
A/Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2240 <SEP>
A/Cross-references: EMBL:AL109747; PID:CA852354.1; GSPDB:GNO0070; SCODEB:SCJ21.08
A/Experimental source: strain A3(2)
C/Genetics:
A:Gene: SCODEB:SCJ21.08
C/Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology
F:24-435/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS>

Query Match

Best Local Similarity 100.0%; Pred. No. 76;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 PAAVPVP 15

Db 1414 PAAVPVP 1420

Search completed: January 14, 2003, 17:14:59
Job time: 7.6004 secs


```

FT Misc-difference 64 /note= "Preferably Leu"
FT Misc-difference 73 /note= "Preferably Asn"
FT MO200044403-A1.
PD 03-AUG-2000.
XX 20-JAN-2000; 2000MO-US01484.
XX 20-JAN-1999; 99US-0234208.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Doherty JK, Clinton GM, Adelman JP;
XX WPI: 2000-499287/44.
DR
XX Using polypeptides and antibodies that bind to the extracellular domain
PT of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the
PT breast, lung, ovaries and colon
XX
PS Claim 1, Page 39; 46pp; English.
XX
CC HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The
CC extracellular domain of p185-HER-2 is proteolytically shed from breast
CC carcinoma cells in culture and is found in serum of some cancer patients
CC and may be a serum marker of metastatic breast cancer. An alternative
CC HER-2 mRNA of 4.8 kb with a 274 bp insert (intron 8) has been
CC identified. The retained intron is in-frame and encodes a 79 amino acid
CC extension designated ECDIIIA (the present sequence), which is inserted at
CC residue 340 of p185-HER-2. The alternative mRNA predicts a truncated
CC HER-2 protein (approximately 68 kDa) that lacks the transmembrane and
CC intracellular domains (see AAY97240.). p68HER-2 specifically binds to
CC p185-HER-2 without activating HER-2. It could therefore block
CC dimerization of p185-HER-2. The p68HER-2 polypeptide binds to a site on
CC the ECD of HER-2 that is different from the site of binding for
CC Herceptin (RTM) (a marketed humanized monoclonal antibody that is used
CC for the treatment of cancer and binds to the ECD of HER-2). The methods,
CC compositions, polypeptides and antibodies are used to treat solid
CC tumors such as breast cancer, small cell lung carcinoma, ovarian cancer
CC and/or colon cancer, especially where over-expression of HER-2 is
CC indicated.
XX
SQ Sequence 79 AA:
Query Match 87.3%; Score 69; DB 21; Length 79;
Best Local Similarity 100.0%; Pred. No. 3.2e-66;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GXHSXPRPAVVPYRXQPRPAHPLYSLFRLPSMDYVSAFYSLPLAPLSPTSYXISPSV 60
DB 1 GXHSXPRPAVVPYRXQPRPAHPLYSLFRLPSMDYVSAFYSLPLAPLSPTSYXISPSV 60
QY 61 GRGXDPDAHVAVXLSRYEG 79
DB 61 GRGXDPDAHVAVXLSRYEG 79
RESULT 2
AAE09180
ID AAE09180 standard; peptide; 79 AA.
XX
XX AAE09180;
AC
XX 15-NOV-2001 (first entry)
DT
XX Human p68HER-2 ECDIIIA generic sequence #1.
DE
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KM solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA.

```

```

XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Misc-difference 2 /label= Unknown
FT /note= "Encoded by WCC"
FT Misc-difference 5 /label= Unknown
FT /note= "Encoded by WCC"
FT Misc-difference 6 /label= Unknown
FT /note= "Encoded by CYC"
FT Misc-difference 16 /label= Unknown
FT /note= "Encoded by CYC"
FT Misc-difference 18 /label= Unknown
FT /note= "Encoded by CWC"
FT Misc-difference 21 /label= Unknown
FT /note= "Encoded by AYR"
FT Misc-difference 36 /label= Unknown
FT /note= "Encoded by GNC"
FT Misc-difference 54 /label= Unknown
FT /note= "Encoded by MTA"
FT Misc-difference 64 /label= Unknown
FT /note= "Encoded by CSR"
FT Misc-difference 73 /label= Unknown
FT /note= "Encoded by CYG"
FT Misc-difference 73 /label= Unknown
FT /note= "Encoded by SAC"
XX WO200161356-A1.
XX
XX 23-AUG-2001.
XX
XX 16-FEB-2001; 2001MO-US05327.
XX
XX 16-FEB-2000; 2000US-0506079.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX
XX Clinton G, Hennen WD, Evans A;
XX WPI: 2001-529934/58.
XX N-PSDB; AAD15844.
XX
XX New polypeptide, which binds to the extracellular domain of HER-2 for
XX the treatment of hard tumors -
XX
PS Claim 1, Page 52-53; 61pp; English.
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
XX particularly a polypeptide that binds to the extracellular domain (ECD)
XX of HER-2 at a site that is different from the binding site of humanised
XX antibody, Herceptin, at an affinity of at least 108. The present
XX invention is based upon the initial discovery of an alternative HER-2
XX mRNA transcript with 274 bp insert of intron 8. The translation product
XX of the alternative transcript is a truncated HER-2 protein designated
XX p68HER-2 which lacks the transmembrane and intracellular domains of
XX p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
XX The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
XX the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX solid tumours. The present sequence is human p68HER-2 ECDIIIA
XX peptide generic sequence.
XX
SQ Sequence 79 AA:
Query Match 87.3%; Score 69; DB 22; Length 79;

```



```

FT FT /label= Unknown
FT FT /note= "Encoded by GNC"
FT FT Misc-difference 31
FT FT /label= Unknown
FT FT /note= "Encoded by AKA"
FT FT Misc-difference 36
FT FT /label= Unknown
FT FT /note= "Encoded by MTA"
FT FT Misc-difference 54
FT FT /label= Unknown
FT FT /note= "Encoded by CSR"
FT FT Misc-difference 64
FT FT /label= Unknown
FT FT /note= "Encoded by CYG"
FT FT Misc-difference 73
FT FT /label= Unknown
FT FT /note= "Encoded by VAC"
XX PN WO200161356-A1.
XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US05327.
XX PR 16-FEB-2000; 2000US-0506079.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G, Hennen WD, Evans A;
XX DR WPI: 2001-529934/58.
XX DR N-PSDB; AADI5852.
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for
XX PT the treatment of hard tumors -
XX PS Claim 1: Page 57; 61pp; English.
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist
XX CC particularly a polypeptide that binds to the extracellular domain (ECD)
XX CC of HER-2 at a site that is different from the binding site of humanised
XX CC antibody, Herceptin, at an affinity of at least 10-8. The present
XX CC invention is based upon the initial discovery of an alternative HER-2
XX CC mRNA transcript with 274 bp insert of intron 8. The translation product
XX CC of the alternative transcript is a truncated HER-2 protein designated
XX CC p68HER-2 which lacks the transmembrane and intracellular domains of
XX CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIia.
XX CC The ECDIIia-containing polypeptides bind tightly to, and thus antagonise
XX CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX CC nucleic acids encoding these are useful to treat, diagnose and identify
XX CC solid tumours. The present sequence is human p68HER-2 ECDIIia
XX CC peptide generic sequence.
SQ Sequence 79 AA;
Query Match 55.7%; Score 44; DB 22; Length 79;
Best local Similarity 100.0%; Pred. No. 1.9e-39;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 32 PSMDEXSAFSLPLAPLPTSVXISPVSGRGKDPDAHVAVXLSRREG 79
DB 32 PSMDEXSAFSLPLAPLPTSVXISPVSGRGKDPDAHVAVXLSRREG 79
RESULT 5
AAE09181
ID AAE09181 standard; Protein; 419 AA.
XX AC AAE09181;
XX DT 15-NOV-2001 (first entry)
XX DE Human p68HER-2 generic sequence #1.

```

```

XX XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
XX KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
XX KM p68HER-2; ECDIIia.
XX OS Homo sapiens.
XX FH Key
XX FH Region
XX FT 1..340
XX FT /note= "Identical to N-terminal region of p185HER-2"
XX FT Domain
XX FT 341..419
XX FT /label= ECDIIia
XX FT /note= "Extracellular domain IIIa"
XX FT Misc-difference 124
XX FT /note= "Represented as Agn in the sequence shown in
XX FT the specification"
XX FT Misc-difference 125
XX FT /note= "Represented as Agn in the sequence shown in
XX FT the specification"
XX FT Misc-difference 342
XX FT /label= Unknown
XX FT /note= "Encoded by MCC"
XX FT Misc-difference 345
XX FT /label= Unknown
XX FT /note= "Encoded by CYG"
XX FT Misc-difference 346
XX FT /label= Unknown
XX FT /note= "Encoded by CYC"
XX FT Misc-difference 356
XX FT /label= Unknown
XX FT /note= "Encoded by CWG"
XX FT Misc-difference 358
XX FT /label= Unknown
XX FT /note= "Encoded by ATP"
XX FT Misc-difference 361
XX FT /label= Unknown
XX FT /note= "Encoded by GNC"
XX FT Misc-difference 376
XX FT /label= Unknown
XX FT /note= "Encoded by MTA"
XX FT Misc-difference 389
XX FT /note= "Encoded by AGC"
XX FT Misc-difference 394
XX FT /label= Unknown
XX FT /note= "Encoded by CSR"
XX FT Misc-difference 404
XX FT /label= Unknown
XX FT /note= "Encoded by CYG"
XX FT Misc-difference 413
XX FT /label= Unknown
XX FT /note= "Encoded by SAC"
XX PN WO200161356-A1.
XX PD 23-AUG-2001.
XX PF 16-FEB-2001; 2001WO-US05327.
XX PR 16-FEB-2000; 2000US-0506079.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Clinton G, Hennen WD, Evans A;
XX DR WPI: 2001-529934/58.
XX DR N-PSDB; AADI5844.
XX PT New polypeptide, which binds to the extracellular domain of HER-2 for
XX PT the treatment of hard tumors -
XX PS Claim 8: Page 53-54; 61pp; English.
XX CC The invention relates to novel HER-2 (herstatin-2) antagonist

```

CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 protein
CC containing ECDIIIA generic sequence.
XX
XX
SQ Sequence 419 AA;

Query Match 51.9%; Score 41; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.5e-35;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GXHSXPRPAVPVYRXQXPAPHPVLSFLRPSMDXVSATFSLPLAPL 48
Db 341 GXHSXPRPAVPVYRXQXPAPHPVLSFLRPSMDXVSATFSLPLAPL 368
|||||
RESULT 6
AAE09208 ID AAE09208 standard; Protein: 419 AA.
XX
AC AAE09208;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 6.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Domain 341..419
FT /label= ECDIIIA, variant
FT /note= "Extracellular domain IIIA variant"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 125
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 342
FT /note= "Represented as Agn in the parent sequence shown
FT in the specification"
FT Misc-difference 345
FT /label= Unknown
FT Misc-difference 346
FT /label= Unknown
FT Misc-difference 356
FT /label= Unknown
FT Misc-difference 358
FT /label= Unknown
FT Misc-difference 361
FT /label= Unknown
FT Misc-difference 376
FT /label= Asp, Ala, Val
FT Misc-difference 394
FT /label= Unknown
FT Misc-difference 404
FT /label= Unknown
FT Misc-difference 413
FT /label= Unknown

XX
PN W0200161356-A1.
XX
XX 23-AUG-2001.
PD
XX 16-FEB-2001; 2001MO-US05327.
PE
XX 16-FEB-2000; 2000US-0506079.
PR
XX 16-FEB-2000; 2000US-0506079.
XX
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Clinton G, Hemmer WD, Evans A;
XX
XX WPI; 2001-529934/58.
DR
XX
XX
PT New polypeptide, which binds to the extracellular domain of HER-2 for
PT the treatment of hard tumors -
XX
XX Example 11; Page -; 61pp; English.
PS
XX
XX The invention relates to novel HER-2 (herstatin-2) antagonist
CC particularly a polypeptide that binds to the extracellular domain (ECD)
CC of HER-2 at a site that is different from the binding site of humanised
CC antibody, Herceptin, at an affinity of at least 10⁻⁸. The present
CC invention is based upon the initial discovery of an alternative HER-2
CC mRNA transcript with 274 bp insert of intron 8. The translation product
CC of the alternative transcript is a truncated HER-2 protein designated
CC p68HER-2 which lacks the transmembrane and intracellular domains of
CC p185HER-2 but contains ECD 1, II of the p185HER-2 and the novel ECDIIIA.
CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC nucleic acids encoding these are useful to treat, diagnose and identify
CC solid tumours. The present sequence is human p68HER-2 generic protein
CC containing ECDIIIA variant sequence.
CC Note: The present sequence is not shown in the specification but is
CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
CC sequence listing (AAE09181).
XX
XX
SQ Sequence 419 AA;

Query Match 51.9%; Score 41; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.5e-35;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GXHSXPRPAVPVYRXQXPAPHPVLSFLRPSMDXVSATFSLPLAPL 48
Db 341 GXHSXPRPAVPVYRXQXPAPHPVLSFLRPSMDXVSATFSLPLAPL 368
|||||
RESULT 7
AAE09210 ID AAE09210 standard; Protein: 419 AA.
XX
AC AAE09210;
XX
DT 15-NOV-2001 (first entry)
XX
DE Human p68HER-2 generic protein variant 8.
XX
XX HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KM p68HER-2; ECDIIIA; variant.
XX
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..340
FT /note= "Identical to N-terminal region of p185HER-2"
FT Domain 341..419
FT /label= ECDIIIA, variant
FT /note= "Extracellular domain IIIA variant"
FT Misc-difference 124
FT /note= "Represented as Agn in the parent sequence shown

```

FT      in the specification"
FT      Misc-difference 125
FT      /note= "Represented as Agn in the parent sequence shown
FT      in the specification"
FT      Misc-difference 342
FT      /label= Unknown
FT      Misc-difference 345
FT      /label= Unknown
FT      Misc-difference 346
FT      /label= Unknown
FT      Misc-difference 356
FT      /label= Unknown
FT      Misc-difference 358
FT      /label= Unknown
FT      Misc-difference 361
FT      /label= Unknown
FT      Misc-difference 376
FT      /label= Unknown
FT      Misc-difference 394
FT      /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT      substituted with Arg"
FT      Misc-difference 404
FT      /label= Unknown
FT      Misc-difference 413
FT      /label= Unknown
FT      WO200161356-A1.
FT      23-AUG-2001.
FT      16-FEB-2001: 2001WO-US05327.
FT      16-FEB-2000: 2000US-0506079.
FT      (UYOR-) UNIV OREGON HEALTH SCI.
FT      Clinton G, Hennen WD, Evans A;
FT      WPI; 2001-529934/58.
FT      New polypeptide, which binds to the extracellular domain of HER-2 for
FT      the treatment of hard tumors -
FT      Example 11; Page -: 61pp; English.
FT      The invention relates to novel HER-2 (herstatin-2) antagonist
FT      particularly a polypeptide that binds to the extracellular domain (ECD)
FT      of HER-2 at a site that is different from the binding site of humanised
FT      antibody, Herceptin, at an affinity of at least 108. The present
FT      invention is based upon the initial discovery of an alternative HER-2
FT      mRNA transcript with 274 bp insert of intron 8. The translation product
FT      of the alternative transcript is a truncated HER-2 protein designated
FT      p68HER-2 which lacks the transmembrane and intracellular domains of
FT      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
FT      The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
FT      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
FT      nucleic acids encoding these are useful to treat, diagnose and identify
FT      solid tumours. The present sequence is human p68HER-2 generic protein
FT      containing ECDIIIA variant sequence.
FT      Note: The present sequence is not shown in the specification but is
FT      derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
FT      sequence listing (AAE09181).
FT      SQ      Sequence      419 AA:
FT      Query Match      51.9%: Score 41; DB 22; Length 419;
FT      Best Local Similarity 100.0%; Pred.No.1.5e-35;
FT      Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT      1 GXHSXXRPAVAVPXPXQXPXAHNPVLSFLRPSMDXVSATFSLAPL 48
FT      |||||||
FT      341 GXHSXXRPAVAVPXPXQXPXAHNPVLSFLRPSMDXVSATFSLAPL 388

```

```

RESULT 8
AAE09211
ID      AAE09211 standard; Protein; 419 AA.
XX
AC      AAE09211;
XX
DE      15-NOV-2001 (first entry)
XX
DE      Human p68HER-2 generic protein variant 9.
XX
KW      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW      solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW      p68HER-2; ECDIIIA; variant.
XX
OS      Homo sapiens.
XX
FH      Key
FH      Region
FT      1..340
FT      /note= "Identical to N-terminal region of p185HER-2"
FT      341..419
FT      /label= ECDIIIA variant
FT      /note= "Extracellular domain IIIa variant"
FT      Misc-difference 124
FT      /note= "Represented as Agn in the parent sequence shown
FT      in the specification"
FT      Misc-difference 125
FT      /note= "Represented as Agn in the parent sequence shown
FT      in the specification"
FT      Misc-difference 342
FT      /label= Unknown
FT      Misc-difference 345
FT      /label= Unknown
FT      Misc-difference 346
FT      /label= Unknown
FT      Misc-difference 356
FT      /label= Unknown
FT      Misc-difference 358
FT      /label= Unknown
FT      Misc-difference 361
FT      /label= Unknown
FT      Misc-difference 376
FT      /label= Unknown
FT      Misc-difference 394
FT      /label= Unknown
FT      Misc-difference 404
FT      /label= Unknown
FT      /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT      substituted with Ieu"
FT      Misc-difference 413
FT      /label= Unknown
FT      WO200161356-A1.
FT      23-AUG-2001.
FT      16-FEB-2001: 2001WO-US05327.
FT      16-FEB-2000: 2000US-0506079.
FT      (UYOR-) UNIV OREGON HEALTH SCI.
FT      Clinton G, Hennen WD, Evans A;
FT      WPI; 2001-529934/58.
FT      New polypeptide, which binds to the extracellular domain of HER-2 for
FT      the treatment of hard tumors -
FT      Example 11; Page -: 61pp; English.
FT      The invention relates to novel HER-2 (herstatin-2) antagonist
FT      particularly a polypeptide that binds to the extracellular domain (ECD)
FT      of HER-2 at a site that is different from the binding site of humanised

```

CC	antibody Herceptin, at an affinity of at least 10 ⁻⁸ . The present
CC	invention is based upon the initial discovery of an alternative HER-2
CC	mRNA transcript with 274 bp insert of Intron 8. The translation product
CC	of the alternative transcript is a truncated HER-2 protein designated
CC	p68HER-2 which lacks the transmembrane and intracellular domains of
CC	p15HER-2 but contains ECD 1, II of the p15HER-2 and the novel ECDIIa.
CC	The ECDIIa-containing polypeptides bind tightly to, and thus antagonise
CC	the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC	nucleic acids encoding these are useful to treat, diagnose and identify
CC	solid tumours. The present sequence is human p68HER-2 generic protein
CC	containing ECDIIa variant sequence.
CC	Note: The present sequence is not shown in the specification but is
CC	derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
CC	sequence listing (AAE09181).
XX	
SO	Sequence 419 AA;
QY	1 GXHXXPRPAVPVPRXQXPAPHVLSFLRSPMDXVSASFYSIPLAPL 48
Db	341 GXHXXPRPAVPVPRXQXPAPHVLSFLRSPMDXVSASFYSIPLAPL 388
RESULT 9	
ID	AAE09212
XX	AAE09212 standard; Protein; 419 AA.
XX	AAE09212;
XX	
DT	15-NOV-2001 (first entry)
XX	
DE	Human p68HER-2 generic protein variant 10.
KW	HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
KW	solid tumour; cancer; polymorphism; cytostatic; gene therapy;
KW	p68HER-2; ECDIIa; variant.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Region
FT	Location/Qualifiers
FT	1..340
FT	/note= "Identical to N-terminal region of p15HER-2"
FT	341..419
FT	/label= ECDIIa_variant
FT	/note= "Extracellular domain I1a variant"
FT	124
FT	/note= "Represented as Agn in the parent sequence shown
FT	in the specification"
FT	125
FT	/note= "Represented as Agn in the parent sequence shown
FT	in the specification"
FT	342
FT	/label= Unknown
FT	345
FT	/label= Unknown
FT	346
FT	/label= Unknown
FT	356
FT	/label= Unknown
FT	358
FT	/label= Unknown
FT	361
FT	/label= Unknown
FT	376
FT	/label= Unknown
FT	394
FT	/label= Unknown
FT	404
FT	/label= Unknown
FT	413
FT	/label= Unknown

	/note= "p68HER-2 generic sequence (AAE09181) Xaa substituted with Asn"
FT	WO200161356-A1.
PV	
NN	23 -AUG- 2001.
PD	
XX	16-FEB- 2001; 2001WO-US05327.
PF	
PR	16-FEB- 2000; 2000US-0506079.
XX	(UYOR-) UNIV OREGON HEALTH SCI.
PA	
JT	Clinton G, Henner WD, Evans A;
DR	WPI; 2001-529934/58.
XX	New polypeptide, which binds to the extracellular domain of HER-2 for
PT	the treatment of hard tumors -
PS	Example 11; Page -: 61pp; English.
CC	The invention relates to novel HER-2 (herstatin-2) antagonist
CC	particularly a polypeptide that binds to the extracellular domain (ECD)
CC	of HER-2 at a site that is different from the binding site of humanised
CC	antibody, Herceptin, at an affinity of at least 10 ⁻⁸ . The present
CC	invention is based upon the initial discovery of an alternative HER-2
CC	mRNA transcript with 274 bp insert of intron 8. The translation product
CC	of the alternative transcript is a truncated HER-2 protein designated
CC	p68HER-2 which lacks the transmembrane and intracellular domains of
CC	p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
CC	The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
CC	the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
CC	nucleic acids encoding these are useful to treat, diagnose and identify
CC	solid tumours. The present sequence is human p68HR-2 generic protein
CC	containing ECDIIIA variant sequence.
CC	Note: The present sequence is not shown in the specification but is
CC	derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
SQ	sequence listing (AAE09181).
Sequence	419 AA:
Query Match	51.9%; Score 41; DB 22; Length 419;
Best Local Similarity	100.0%; Pred. No. 1.5e-35;
Matches	48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GXHXXXXPPRANVPVRXKQPAPAHNYLSLRLRSMDXSVAFYSLPLAPL 48
Dd	341 GXHSXXXPRPAAVPVEXRXQXPAPHNVLSPFLRPMSDYSVASFYLPLAPL 388
RESULT 10	
AAE20348	
ID	AAE20348 standard; Protein; 419 AA.
AC	AEE20348;
DT	18-JUN-2002 (first entry)
DE	Human truncated HER2 protein.
KM	Human; tumour; endothelial growth factor receptor; EGFR; cytosolic; herstatin; HER-2 receptor tyrosine kinase; squamous cell carcinoma; lung; colon; glial cell tumour; cell growth.
OS	Homo sapiens.
FH	Key Location/Qualifiers
FT	Misc-difference 342 /label= Thr, Ser
FT	Misc-difference 345 /label= Leu, Pro
FT	Misc-difference 346

FT Misc-difference 356 /label= Pro, Leu
 FT /label= Leu, Gln
 FT Misc-difference 358 /label= Met, Leu
 FT Misc-difference 361 /label= Gly, Asp, Ala, Val
 FT Misc-difference 376 /label= Leu, Ile
 FT Misc-difference 394 /label= Pro, Arg
 FT Misc-difference 404 /label= Pro, Leu
 FT Misc-difference 413 /label= Asp, Asn
 FT WO200214470-A2.
 XX
 PN 21-FEB-2002.
 XX
 PD 14-AUG-2001; 2001WO-US25502.
 XX
 PF 14-AUG-2000; 2000US-0638834.
 XX
 PR (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PA Clinton GM;
 XX
 PI WPI; 2002-269185/31.
 XX
 DR Treating solid tumor characterized by expression of endothelial growth
 XX factor receptor, involves administering recombinant herstatin that
 PT binds to extracellular domain of the endothelial growth factor receptor
 PT
 PT Claim 1; Page 78-80; 82pp; English.
 XX
 PS The present invention relates to a method for treating a solid tumour
 CC characterized by endothelial growth factor receptor (EGFR) expression.
 CC The method involves administering an agent that binds to an extracellular
 CC domain (ECD) of EGFR. The invention also relates to a naturally occurring
 CC inhibitor of HER-2 receptor tyrosine kinase called herstatin. The co-
 CC expression of herstatin with p185HER2 causes a striking reduction in cell
 CC growth that corresponds with suppression of p185 autophosphorylation. The
 CC method or a pharmaceutical composition is useful for treating a solid
 CC tumour (selected from squamous cell carcinoma, lung carcinoma, colon
 CC carcinoma and glioma cell tumour) characterized by EGFR expression. The
 CC present sequence is human truncated HER2 protein that lacks transmembrane
 CC and intracellular domains.
 XX
 SQ Sequence 419 AA;
 Query Match 51.9%; Score 41; DB 23; Length 419;
 Best local Similarity 100.0%; Pred. No. 1.5e-35;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GHSXXPRPAVVPKXQXPRAHPVLSFLRPSMDXVSATYSPLAPL 48
 DB 341 GHSXXPRPAVVPKXQXPRAHPVLSFLRPSMDXVSATYSPLAPL 388
 RESULT 11
 AAAY97240
 ID AAAY97240 standard; protein; 420 AA.
 XX
 AC AAAY97240;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE Truncated HER-2, p68-HER-2.
 XX
 KW HER-2; erbB-2; oncogene; receptor-like tyrosine kinase; insertion;
 extracellular domain IIIa; antagonist; intron 8; C-terminal extension;

KW truncated HER-2; p68; dimerization inhibitor; cytostatic.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 124 /note= "represented as Agn"
 FT Misc-difference 125 /note= "represented as Agn"
 FT Misc-difference 125 /note= "represented as Agn"
 FT Misc-difference 143 /note= "Preferably Ser"
 FT Misc-difference 346 /note= "Preferably Pro"
 FT Misc-difference 347 /note= "Preferably Leu"
 FT Misc-difference 357 /note= "Preferably Gln"
 FT Misc-difference 359 /note= "Preferably Leu"
 FT Misc-difference 362 /note= "Changes from glycine"
 FT Misc-difference 377 /note= "Preferably Ile"
 FT Misc-difference 395 /note= "Preferably Arg"
 FT Misc-difference 405 /note= "Preferably Leu"
 FT Misc-difference 414 /note= "Preferably Asn"
 FT WO200044403-A1.
 XX
 PN 03-AUG-2000.
 XX
 PD 20-JAN-2000; 2000WO-US01484.
 XX
 PF 20-JAN-1999; 99US-0234208.
 XX
 PR (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PA Doherty JK, Clinton GM, Adelman JP;
 XX
 PI WPI; 2000-499287/44.
 XX
 DR Using polypeptides and antibodies that bind to the extracellular domain
 XX of the receptor-like tyrosine kinase HER-2 to treat solid tumors of the
 XX breast, lung, ovaries and colon
 PT
 PT Claim 8; Page 39-40; 46pp; English.
 PS
 XX
 CC HER-2/neu (erbB-2) oncogene encodes a receptor-like tyrosine kinase. The
 CC extracellular domain of p185-HER-2 is proteolytically shed from breast
 CC carcinoma cells in culture and is found in serum of some cancer patients
 CC and may be a serum marker of metastatic breast cancer. An alternative
 CC HER-2 mRNA of 4.8 kb with a 274 bp insert (intron 8) has been
 CC identified. The retained intron is in-frame and encodes a 79 amino acid
 CC residue designated ECDIIa (the present sequence), which is inserted at
 CC position 340 of p185-HER-2. The alternative mRNA predicts a truncated
 CC HER-2 protein (approximately 68 kDa) that lacks the transmembrane and
 CC intracellular domains (see AAAY97240). p68HER-2 specifically binds to
 CC p185-HER-2 without activating HER-2. It could therefore block
 CC dimerization of p185-HER-2. The p68HER-2 polypeptide binds to a site on
 CC the ECD of HER-2 that is different from the site of binding for
 CC Herceptin (RTM) (a marketed humanized monoclonal antibody that is used
 CC for the treatment of cancer and binds to the ECD of HER-2). The methods,
 CC compositions, polypeptides and antibodies are used to treat solid
 CC tumours such as breast cancer, small cell lung carcinoma, ovarian cancer
 CC and/or colon cancer, especially where over-expression of HER-2 is
 CC indicated.
 XX
 SQ Sequence 420 AA;
 Query Match 51.9%; Score 41; DB 21; Length 420;

Best Local Similarity 100.0%; Pred. No. 1.5e-35; Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GXHSXPRPAVPYVXKQXPAPHPVLSFLRPSMDXVSATYSLPLAPL 48
 DB 342 GXHSXPRPAVPYVXKQXPAPHPVLSFLRPSMDXVSATYSLPLAPL 369

RESULT 12
 AAE09203
 ID AAE09203 standard; Protein: 419 AA.

AC AAE09203;

DT 15-NOV-2001 (first entry)

DE Human p68HER-2 generic protein variant 1.

KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDIIIA; variant.

OS Homo sapiens.

Key Location/Qualifiers

FT Region 1..340 /note= "Identical to N-terminal region of p185HER-2"

FT Domain 341..419 /label= "ECDIIIA-variant"

FT /note= "Extracellular domain IIIA variant"

FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown in the specification"

FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown in the specification"

FT Misc-difference 342 /note= "p68HER-2 generic sequence (AAE09181) Xaa substituted with Ser"

FT Misc-difference 345

FT /label= Unknown

FT Misc-difference 346

FT /label= Unknown

FT Misc-difference 356

FT /label= Unknown

FT Misc-difference 358

FT /label= Unknown

FT Misc-difference 361

FT /label= Unknown

FT Misc-difference 376

FT /label= Unknown

FT Misc-difference 394

FT /label= Unknown

FT Misc-difference 404

FT /label= Unknown

FT Misc-difference 413

FT /label= Unknown

XX MO200161356-A1.

XX 23-AUG-2001.

XX 16-FEB-2001; 2001MO-US05327.

XX 16-FEB-2000; 2000US-0506079.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Clinton G, Henner WD, Evans A;

XX WPI; 2001-529934/58.

XX New polypeptide, which binds to the extracellular domain of HER-2 for

FT the treatment of hard tumors -
 PS Example 11; Page -: 61pp; English.

CC The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁶. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of Intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDIIIA variant sequence.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
 CC sequence listing (AAE09181).

SO Sequence 419 AA;

Query Match 50.6%; Score 40; DB 22; Length 419;
 Best Local Similarity 100.0%; Pred. No. 1.8e-34; Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 HSXPRPAVPYVXKQXPAPHPVLSFLRPSMDXVSATYSLPLAPL 48
 DB 343 HSXPRPAVPYVXKQXPAPHPVLSFLRPSMDXVSATYSLPLAPL 368

RESULT 13

ID AAE09204 standard; Protein: 419 AA.

AC AAE09204;

DT 15-NOV-2001 (first entry)

DE Human p68HER-2 generic protein variant 2.

KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDIIIA; variant.

OS Homo sapiens.

Key Location/Qualifiers

FT Region 1..340 /note= "Identical to N-terminal region of p185HER-2"

FT Domain 341..419 /label= "ECDIIIA-variant"

FT /note= "Extracellular domain IIIA variant"

FT Misc-difference 124 /note= "Represented as Agn in the parent sequence shown in the specification"

FT Misc-difference 125 /note= "Represented as Agn in the parent sequence shown in the specification"

FT Misc-difference 342

FT /label= Unknown

FT Misc-difference 345 /note= "p68HER-2 generic sequence (AAE09181) Xaa substituted with Pro"

FT Misc-difference 346

FT /label= Unknown

FT Misc-difference 356

FT /label= Unknown

FT Misc-difference 358

FT /label= Unknown

FT Misc-difference 361

```

FT      /label= Unknown
FT      Misc-difference 376
FT      /label= Unknown
FT      Misc-difference 394
FT      /label= Unknown
FT      Misc-difference 404
FT      /label= Unknown
FT      Misc-difference 413
FT      /label= Unknown
FT      WO200161356-A1.
XX      PD      23-AUG-2001.
XX      PF      16-FEB-2001; 2001WO-US05327.
XX      PR      16-FEB-2000; 2000US-0506079.
XX      PA      (UYOR-) UNIV OREGON HEALTH SCI.
XX      PI      Clinton G, Henner WD, Evans A;
XX      DR      WPI; 2001-529934/58.
XX      PT      New polypeptide, which binds to the extracellular domain of HER-2 for
XX      PT      the treatment of hard tumors -
XX      PS      Example 11; Page -: 61pp; English.
XX      CC      The invention relates to novel HER-2 (herstatin-2) antagonist
XX      CC      particularly a polypeptide that binds to the extracellular domain (ECD)
XX      CC      of HER-2 at a site that is different from the binding site of humanised
XX      CC      antibody, Herceptin, at an affinity of at least 10-8. The present
XX      CC      invention is based upon the initial discovery of an alternative HER-2
XX      CC      mRNA transcript with 274 bp insert of intron 8. The translation product
XX      CC      of the alternative transcript is a truncated HER-2 protein designated
XX      CC      p68HER-2 which lacks the transmembrane and intracellular domains of
XX      CC      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIIa.
XX      CC      The ECDIIIIa-containing polypeptides bind tightly to, and thus antagonise
XX      CC      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      CC      nucleic acids encoding these are useful to treat, diagnose and identify
XX      CC      solid tumours. The present sequence is human p68HER-2 generic protein
XX      CC      containing ECDIIIIa variant sequence.
XX      CC      Note: The present sequence is not shown in the specification but is
XX      CC      derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX      CC      sequence listing (AAE09181).
XX      SQ      Sequence 419 AA;
OY      7 PRPAAVPVRKXOPRAPHVPLSRPSMDXVSFYSPLAPL 48
DB      347 PRPAAVPVRKXOPRAPHVPLSRPSMDXVSFYSPLAPL 388
      Query Match      48.1%; Score 38; DB 22; Length 419;
      Best Local Similarity 100.0%; Pred. No. 2.4e-32;
      Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
      RESULT 14
      AAE09205
      ID      AAE09205 standard; Protein; 419 AA.
      XX      AAE09205;
      DE      15-NOV-2001 (first entry)
      XX      Human p68HER-2 generic protein variant 3.
      XX      HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
      KW      solid tumour; cancer; polymorphism; cytostatic; gene therapy;
      KW      p68HER-2; ECDIIIIa; variant.
      XX      OS      Homo sapiens.

```

```

XX      Key      Location/Qualifiers
FH      Region      1..340
FT      Domain      341..419
FT      /note= "Identical to N-terminal region of p185HER-2"
FT      /label= ECDIIIIa_variant
FT      /note= "Extracellular domain IIIa variant"
FT      Misc-difference 124
FT      /note= "Represented as Agn in the parent sequence shown
FT      in the specification"
FT      Misc-difference 125
FT      /note= "Represented as Agn in the parent sequence shown
FT      in the specification"
FT      Misc-difference 342
FT      /label= Unknown
FT      Misc-difference 345
FT      /label= Unknown
FT      Misc-difference 346
FT      /note= "p68HER-2 generic sequence (AAE09181) Xaa
FT      substituted with Leu"
FT      Misc-difference 356
FT      /label= Unknown
FT      Misc-difference 358
FT      /label= Unknown
FT      Misc-difference 361
FT      /label= Unknown
FT      Misc-difference 376
FT      /label= Unknown
FT      Misc-difference 394
FT      /label= Unknown
FT      Misc-difference 404
FT      /label= Unknown
FT      Misc-difference 413
FT      /label= Unknown
XX      WO200161356-A1.
XX      PD      23-AUG-2001.
XX      PF      16-FEB-2001; 2001WO-US05327.
XX      PR      16-FEB-2000; 2000US-0506079.
XX      PA      (UYOR-) UNIV OREGON HEALTH SCI.
XX      PI      Clinton G, Henner WD, Evans A;
XX      DR      WPI; 2001-529934/58.
XX      PT      New polypeptide, which binds to the extracellular domain of HER-2 for
XX      PT      the treatment of hard tumors -
XX      PS      Example 11; Page -: 61pp; English.
XX      CC      The invention relates to novel HER-2 (herstatin-2) antagonist
XX      CC      particularly a polypeptide that binds to the extracellular domain (ECD)
XX      CC      of HER-2 at a site that is different from the binding site of humanised
XX      CC      antibody, Herceptin, at an affinity of at least 10-8. The present
XX      CC      invention is based upon the initial discovery of an alternative HER-2
XX      CC      mRNA transcript with 274 bp insert of intron 8. The translation product
XX      CC      of the alternative transcript is a truncated HER-2 protein designated
XX      CC      p68HER-2 which lacks the transmembrane and intracellular domains of
XX      CC      p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIIa.
XX      CC      The ECDIIIIa-containing polypeptides bind tightly to, and thus antagonise
XX      CC      the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
XX      CC      nucleic acids encoding these are useful to treat, diagnose and identify
XX      CC      solid tumours. The present sequence is human p68HER-2 generic protein
XX      CC      containing ECDIIIIa variant sequence.
XX      CC      Note: The present sequence is not shown in the specification but is
XX      CC      derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
XX      CC      sequence listing (AAE09181).
XX      SQ      Sequence 419 AA;

```


Query Match 48.1%; Score 38; DB 22; Length 419;
 Best Local Similarity 100.0%; Pred. No. 2.4e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 PRPAVPVPPXRXQPPAHPLVLSFLRPSMDXVSAPFSLPLAPL 48
 ||||||||||||||||||||||||||||||||||||||||
 DB 347 PRPAVPVPPXRXQPPAHPLVLSFLRPSMDXVSAPFSLPLAPL 388

RESULT 15
 AAE09213
 ID AAE09213 standard; Protein; 419 AA.
 AC AAE09213;
 DT 15-NOV-2001 (first entry)
 XX
 DE Human p68HER-2 generic protein variant 11.
 XX
 KW HER-2; herstatin; antagonist; extracellular domain; ECD; Herceptin;
 KW solid tumour; cancer; polymorphism; cytostatic; gene therapy;
 KW p68HER-2; ECDIIIA; variant.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FH Region 1..340
 FT /note= "Identical to N-terminal region of p185HER-2"
 FT Domain 341..419
 FT /label= ECDIIIA-variant
 FT /note= "Extracellular domain IIIa variant"
 FT Misc-difference 124
 FT /note= "Represented as Agn in the parent sequence shown
 in the specification"
 FT Misc-difference 125
 FT /note= "Represented as Agn in the parent sequence shown
 in the specification"
 FT Misc-difference 342
 FT /label= Unknown
 FT Misc-difference 345
 FT /label= Unknown
 FT Misc-difference 346
 FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
 substituted with Leu"
 FT Misc-difference 356
 FT /label= Unknown
 FT Misc-difference 358
 FT /label= Unknown
 FT Misc-difference 361
 FT /label= Unknown
 FT Misc-difference 376
 FT /label= Unknown
 FT Misc-difference 394
 FT /label= Unknown
 FT Misc-difference 404
 FT /label= Unknown
 FT Misc-difference 413
 FT /note= "p68HER-2 generic sequence (AAE09181) Xaa
 substituted with Asn"
 FT
 XX WO200161356-A1.
 XX
 XX 23-AUG-2001.
 XX
 XX 16-FEB-2001; 2001WO-US05327.
 XX
 XX 16-FEB-2000; 2000US-0506079.
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 XX Clinton G, Henner WD, Evans A;
 XX

DR WPI: 2001-529934/58.
 XX
 XX New polypeptide, which binds to the extracellular domain of HER-2 for
 PT the treatment of hard tumors -
 PT
 XX
 XX Example 11; Page -: 61pp; English.
 PS
 XX
 XX The invention relates to novel HER-2 (herstatin-2) antagonist
 CC particularly a polypeptide that binds to the extracellular domain (ECD)
 CC of HER-2 at a site that is different from the binding site of humanised
 CC antibody, Herceptin, at an affinity of at least 10⁷8. The present
 CC invention is based upon the initial discovery of an alternative HER-2
 CC mRNA transcript with 274 bp insert of Intron 8. The translation product
 CC of the alternative transcript is a truncated HER-2 protein designated
 CC p68HER-2 which lacks the transmembrane and intracellular domains of
 CC p185HER-2 but contains ECD I, II of the p185HER-2 and the novel ECDIIIA.
 CC The ECDIIIA-containing polypeptides bind tightly to, and thus antagonise
 CC the HER-2 receptor. The peptides, which bind to an HER-2 ECD, and the
 CC nucleic acids encoding these are useful to treat, diagnose and identify
 CC solid tumours. The present sequence is human p68HER-2 generic protein
 CC containing ECDIIIA variant sequence.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from p68HER-2 generic sequence (SEQ ID NO:2) shown in the
 CC sequence listing (AAE09181).
 XX
 SQ Sequence 419 AA;

OY 7 PRPAVPVPPXRXQPPAHPLVLSFLRPSMDXVSAPFSLPLAPL 48
 ||||||||||||||||||||||||||||||||||||||||
 DB 347 PRPAVPVPPXRXQPPAHPLVLSFLRPSMDXVSAPFSLPLAPL 388

Search completed: January 14, 2003, 17:12:54
 Job time : 11.1526 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:10:35 ; Search time 14.3032 Seconds
(without alignments)
1215.012 Million cell updates/sec

Title: US-09-506-079e-2
Perfect score: 419

Sequence: 1 MEALACRNGGLIALLPFGA.....VGRGKDPDAHVAVLSRYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	19.8	1255	1	ERB2_HUMAN
2	51	12.2	1254	1	ERB2_MESAU
3	51	12.2	1257	1	ERB2_RAT
4	9	2.1	230	1	RNFE_VIBCH
5	9	2.1	703	1	EGFR_CHICK
6	9	2.1	1210	1	EGFR_HUMAN
7	9	2.1	1210	1	EGFR_MOUSE
8	8	1.9	276	1	DHCA_RAT
9	8	1.9	463	1	YV30_MCTU
10	8	1.9	725	1	NCA2_MOUSE
11	8	1.9	858	1	NCA1_RAT
12	8	1.9	1115	1	NCA1_MOUSE
13	7	1.7	62	1	SECE_SUISO
14	7	1.7	111	1	MADE_HUMAN
15	7	1.7	114	1	GNZ2_TUPGB
16	7	1.7	173	1	NU6M_CHICK
17	7	1.7	192	1	UREF_ALCEU
18	7	1.7	220	1	GF33_CHICK
19	7	1.7	221	1	SDFL_HUMAN
20	7	1.7	224	1	TCMP_SALT
21	7	1.7	228	1	YU02_HUMAN
22	7	1.7	248	1	DSBG_ECO57
23	7	1.7	248	1	DSBG_ECOLI
24	7	1.7	271	1	HMGL_TRYBR
25	7	1.7	287	1	YCXE_BACSU
26	7	1.7	306	1	SRTI_SYN7
27	7	1.7	307	1	MMPI_MYCLE
28	7	1.7	313	1	ENVI_MOUSE
29	7	1.7	334	1	G3P1_BACSU
30	7	1.7	341	1	XERD_SEIRU
31	7	1.7	357	1	MLTC_HAEIN
32	7	1.7	365	1	YM27_YEAST
33	7	1.7	379	1	IRKA_RAT

34	7	1.7	396	1	YD18_YEAST
35	7	1.7	399	1	BR33-SHEEP
36	7	1.7	400	1	RENI_SHEEP
37	7	1.7	443	1	CG1H_XANMA
38	7	1.7	443	1	SOX3_HUMAN
39	7	1.7	458	1	RADA_HAEIN
40	7	1.7	463	1	MRCO_MESAU
41	7	1.7	524	1	YPBT_BOVIN
42	7	1.7	557	1	YP85_MYCTU
43	7	1.7	574	1	ZEST_DROME
44	7	1.7	575	1	LNK_HUMAN
45	7	1.7	607	1	ND5M_CARAU

ALIGNMENTS

RESULT 1
ERB2_HUMAN STANDARD; PRT; 1255 AA.
AC P04626;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (P185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell
DE surface receptor HER2) (MIM 19).
GN ERB2 OR HER2 OR NGL OR NEU.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86118663; PubMed-3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,
SA Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to
RT epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-86070181; PubMed-2999974;
RA Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,
RA McGarath J., Seeburg P.H., Libermann T.A., Schlessinger J.,
RA Franke J., Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
RN [3]
RP SEQUENCE OF 737-1031 FROM N.A.
RX MEDLINE-86016729; PubMed-2995967;
RA Samba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the
RT c-erbB-1/epidermal growth factor-receptor gene and is amplified in a
RT human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
RN [4]
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE-93194196; PubMed-8095486;
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERB2 gene by allele-
RT specific competition hybridization.";
RL Genomics 15:426-429(1993).
CC - FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
CC ALPHA AND AMPHIREGULIN.
CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC - SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC - SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES (BY SIMILARITY).
 CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN
 CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY
 CC OF 0.782; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;
 CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: M11767; AAA35808.1; JOINED.
 DR EMBL: M11761; AAA35808.1; JOINED.
 DR EMBL: M11762; AAA35808.1; JOINED.
 DR EMBL: M11763; AAA35808.1; JOINED.
 DR EMBL: M11764; AAA35808.1; JOINED.
 DR EMBL: M11765; AAA35808.1; JOINED.
 DR EMBL: M11766; AAA35808.1; JOINED.
 DR EMBL: M11730; AAA35808.1; JOINED.
 DR EMBL: M12036; AAA35978.1; JOINED.
 DR EMBL: X03363; CA27060.1; JOINED.
 DR PIR: A25491; A25491.
 DR PIR: A24571; A24571.
 DR HSSP: P11362; 1FGK.
 DR Genew: HGNC:3430; ERBB2.
 DR MIM: 164870; ERBB2.
 DR InterPro: IPR000494; EGFR_Ldomain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF000659; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_Ldomain; 2.
 DR Pfam: PF02757; YLP; 2.
 DR Prodom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00261; FU; 3.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Polymorphism.
 KW SIGNAL.
 FT CHAIN 1 21 POTENTIAL.
 FT CHAIN 22 1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 653 675 POTENTIAL.
 FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 720 987 PROTEIN KINASE.
 FT NP_BIND 726 734 ATP (BY SIMILARITY).
 FT BINDING 753 753 ATP (BY SIMILARITY).
 FT ACT_SITE 845 845 BY SIMILARITY.
 FT DISULFID 195 204 BY SIMILARITY.
 FT DISULFID 199 212 BY SIMILARITY.
 FT DISULFID 220 227 BY SIMILARITY.
 FT DISULFID 224 225 BY SIMILARITY.
 FT DISULFID 236 244 BY SIMILARITY.
 FT DISULFID 240 252 BY SIMILARITY.
 FT DISULFID 255 264 BY SIMILARITY.
 FT DISULFID 268 295 BY SIMILARITY.
 FT DISULFID 299 311 BY SIMILARITY.
 FT DISULFID 315 331 BY SIMILARITY.
 FT DISULFID 334 338 BY SIMILARITY.
 FT DISULFID 511 520 BY SIMILARITY.
 FT DISULFID 515 528 BY SIMILARITY.
 FT DISULFID 531 540 BY SIMILARITY.
 FT DISULFID 544 560 BY SIMILARITY.

FT DISULFID 563 576 BY SIMILARITY.
 FT DISULFID 567 584 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 600 623 BY SIMILARITY.
 FT DISULFID 626 634 BY SIMILARITY.
 FT DISULFID 630 642 BY SIMILARITY.
 FT MOD_RES 1139 1139 BY SIMILARITY.
 FT MOD_RES 1248 1248 BY SIMILARITY.
 FT CARBOHYD 68 68 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 124 124 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 571 571 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 629 629 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARIANT 654 654 I->V.
 FT VARIANT 655 655 /FTID=VAR_004077.
 FT VARIANT 655 655 I->V.
 FT CONFLICT 1170 1170 /FTID=VAR_004078.
 FT CONFLICT 1170 1170 P->A (IN REF. 2).
 SQ SEQUENCE 1255 AA; 137909 MW; 39E9DFDA04DCF962 CRC64;
 Query Match 19.8%; Score 83; DB 1; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 6.5e-76;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 199 CQSRGWGSESDCQSLRTVCAGCARGCPUPDCCHEQCACCTGPKHSCLACLHF 258
 DB 199 CQSRGWGSESDCQSLRTVCAGCARGCPUPDCCHEQCACCTGPKHSCLACLHF 258
 QY 259 NNSGICELHCPALVYNTDFES 281
 DB 259 NNSGICELHCPALVYNTDFES 281
 RESULT 2
 EMBL_MESAU STANDARD; PRT; 1254 AA.
 ID EMBL_MESAU
 AC 060553;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
 GN ERBB2 OR NEU.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OX NCBI_TaxID=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Nerve;
 RX MEDLINE=94193007; PubMed=7908275;
 RA Nakamura T., Ushijima T., Ishikawa Y., Nagao M., Arai M.,
 RA Yamazaki Y., Ishikawa T.;
 RL "Cloning and activation of the Syrian hamster neu proto-oncogene.";
 RL Gene 140:251-255(1994).
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
 CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```

DR pfam: PF00069; kinase; 1.
DR pfam: PF00757; Furin-like; 1.
DR pfam: PF01030; Recep_L_domain; 2.
DR pfam: PF02757; YLP; 2.
DR pfam: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; FU; 3.
DR SMART: SM00219; TYRKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR Transmembrane: Glycoprotein: Multigene family: Receptor: Signal:
KW Transmembrane: Tyrosine-protein kinase: ATP-binding; Phosphorylation;
KW Proto-oncogene; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 1 21
FT DOMAIN 22 1257
FT TRANSMEM 655 677
FT DOMAIN 678 1257
FT DOMAIN 159 369
FT DOMAIN 473 646
FT DOMAIN 722 989
FT NP_BIND 728 736
FT BINDING 755 755
FT ACT_SITE 847 847
FT DISULFID 196 205
FT DISULFID 200 213
FT DISULFID 221 228
FT DISULFID 225 236
FT DISULFID 237 245
FT DISULFID 241 253
FT DISULFID 256 265
FT DISULFID 269 296
FT DISULFID 300 312
FT DISULFID 316 332
FT DISULFID 335 339
FT DISULFID 513 522
FT DISULFID 517 530
FT DISULFID 533 542
FT DISULFID 546 562
FT DISULFID 565 578
FT DISULFID 589 598
FT DISULFID 602 625
FT DISULFID 628 636
FT DISULFID 632 644
FT MOD_RES 1141 1141
FT MOD_RES 1250 1250
FT CARBOHYD 68 68
FT CARBOHYD 188 188
FT CARBOHYD 260 260
FT CARBOHYD 332 332
FT CARBOHYD 573 573
FT CARBOHYD 631 631
FT VARIANT 661 661
FT V -> E (IN ONCOGENIC NDU).
SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;

Query Match 12.2%; Score 51; DB 1; Length 1257;
Best Local Similarity 100.0%; Pred. No. 2.2e-43;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

GN NFE OR VC1012.
OS Bacteriophage.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EI for N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Yamateyan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.D., Venter J.C.,
RA Fraser C.M.;
RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NQRDE/RNFAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AEO04183; AAF94173.1; -.
DR TIGR: VC1012; -.
DR InterPro: IPR003667; Rnf_Nqr; 1.
DR Pfam: PF02508; Rnf_Nqr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 39 59
FT TRANSMEM 69 89
FT TRANSMEM 93 113
FT TRANSMEM 124 144
FT TRANSMEM 182 202
FT TRANSMEM 24710 24710 MW; 262D4D792044D769 CRC64;
SQ SEQUENCE 230 AA; 24710 MW; 262D4D792044D769 CRC64;

Query Match 2.1%; Score 9; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 4
ID RNF_VIBCH STANDARD; PRT; 230 AA.
AC Q9KT91;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Electron transport complex protein rnfE.

```

```

RESULT 5
ID EGFR_CHICK STANDARD; PRT; 703 AA.
AC P13387;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)
DE (Fragment).
GN EGFR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxId=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88261272; PubMed=3260329;

```


RT isoforms.";
 RL Genomics 71:1-20(2001).
 RN [6]
 RA SEQUENCE OF 575-687 FROM N.A.
 RP Reiter J.L., Thredgill D.W., Danielson A.J., Sehnal C.M.,
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
 RA Mahle N.J.,
 RT "Human and mouse alternative EGFR transcripts encoding only the
 RT extracellular domain of the receptor.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 713-924 FROM N.A.
 RX MEDLINE=84196372; PubMed=6326261;
 RA Lin C.R., Chen W.S., Kruliger W., Stolarsky I.S., Weber W.,
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.,
 RT "Expression cloning of human EGF receptor complementary DNA: gene
 RT amplification and three related messenger RNA products in A431
 RT cells.";
 RL Science 224:843-848(1984).
 RN [8]
 RP SEQUENCE OF 150-962 FROM N.A.
 RX MEDLINE=84245835; PubMed=6330563;
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,
 RA Roe B.A., Merlino G.T., Pastan I.,
 RT "Human epidermal growth factor receptor cDNA is homologous to a
 RT variety of RNAs overproduced in A431 carcinoma cells.";
 RL Nature 309:806-810(1984).
 RN [9]
 RP SEQUENCE OF 1028-1210 FROM N.A.
 RX MEDLINE=85046483; PubMed=6093780;
 RA Stimmgen F.A., Gope M.L., Schulz T.Z., Wright D.A., Carpenter G.,
 RA O'Malley B.W.,
 RT "Isolation of an evolutionarily conserved epidermal growth factor
 RT receptor cDNA from human A431 carcinoma cells.";
 RL Biochem. Biophys. Res. Commun. 124:125-132(1984).
 RN [10]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=88217333; PubMed=3329716;
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,
 RA Waterfield M.D.,
 RT "The human EGF receptor gene: structure of the 110 kb locus and
 RT identification of sequences regulating its transcription.";
 RL Oncogene Res. 1:375-396(1987).
 RN [11]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=91107677; PubMed=1988448;
 RA Haley J.D., Waterfield M.D.,
 RT "Contributory effects of de novo transcription and premature
 RT transcript termination in the regulation of human epidermal growth
 RT factor receptor proto-oncogene RNA synthesis.";
 RL J. Biol. Chem. 266:1746-1753(1991).
 RN [12]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE=85270438; PubMed=2991899;
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.,
 RT "Characterization and sequence of the promoter region of the human
 RT epidermal growth factor receptor gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).
 RN [13]
 RP SEQUENCE OF 540.
 RA Kohda D.,
 RT Submitted (SEP-1997) to the SWISS-PROT data bank.
 RN [14]
 RP RECEPTOR ACTIVITY.
 RX MEDLINE=84191554; PubMed=6325948;
 RA Mroczkowski B., Morig G., Cohen S.,
 RT "ATP-stimulated interaction between epidermal growth factor receptor
 RT and supercoiled DNA.";
 RL Nature 309:270-273(1984).
 RN [15]
 RP PHOSPHORYLATION.
 RX MEDLINE=89278137; PubMed=2543678;
 RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,

RA Howk R., Givol D., Ullrich A., Schlessinger J.,
 RT "All autophosphorylation sites of epidermal growth factor (EGF)
 RT receptor and HER2/neu are located in their carboxyl-terminal tails.
 RT Identification of a novel site in EGF receptor.";
 RL J. Biol. Chem. 264:10667-10671(1989).
 RN [16]
 RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND
 RP ASN-528.
 RX MEDLINE=96398132; PubMed=8962717;
 RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.,
 RT "Analysis of the glycosylation patterns of the extracellular domain of
 RT the epidermal growth factor receptor expressed in Chinese hamster
 RT ovary fibroblasts.";
 RL Growth Factors 13:121-132(1996).
 RN [17]
 RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND
 RP ASN-603.
 RX MEDLINE=20198209; PubMed=1073168;
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.,
 RA "Characterization of the N-oligosaccharides attached to the atypical
 RT Asn-X-Cys sequence of recombinant human epidermal growth factor
 RT receptor.";
 RL J. Biochem. 127:65-72(2000).
 RN [18]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE=98225196; PubMed=9556602;
 RA Abe Y., Otake M., Inagaki F., Lax I., Schlessinger J., Kohda D.,
 RT "Disulfide bond structure of human epidermal growth factor receptor.";
 RL J. Biol. Chem. 273:11150-11157(1998).
 RN [19]
 RP REVIEW.
 RX MEDLINE=87297456; PubMed=3039909;
 RA Carpenter G.,
 RT "Receptors for epidermal growth factor and other polypeptide
 RT mitogens.";
 RL Annu. Rev. Biochem. 56:881-914(1987).
 CC -I- FUNCTION: Receptor for EGF, but also for other members of the EGF
 CC family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding
 CC EGF-like growth factor, GP30 and vaccinia virus growth factor. Is
 CC involved in the control of cell growth and differentiation.
 CC -I- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is
 CC secreted.
 CC -I- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/
 CC truncated isoform/TEGFR, 3/p110 and 4; are produced by
 CC alternative splicing.
 CC -I- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also
 CC expressed in ovarian cancers.
 CC -I- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X00588; CAA25240.1; -
 DR EMBL: U95089; AAB53063.1; -
 DR EMBL: U48722; AAC50802.1; -
 DR EMBL: U48723; AAC50804.1; -
 DR EMBL: U48724; AAC50796.1; -
 DR EMBL: U48725; AAC50797.1; -
 DR EMBL: U48726; AAC50798.1; -

FT	CARBOHYD	175	175	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	352	352	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	413	413	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	444	444	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	528	528	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	568	568	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	603	603	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	623	623	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CONFLICT	19	19	C -> S (IN REF. 2).					
FT	CONFLICT	539	539	C -> W (IN REF. 5).					
FT	CONFLICT	991	991	L -> F (IN REF. 4).					
FT	CONFLICT	1116	1117	HP -> DR (IN REF. 6).					
SO	SEQUENCE	1210 AA;	134853 MW;	690E20D46DFD2DF5 CRC64;					
Query Match		2.1%;	Score 9;	DB 1;	Length 1210;				
Best Local Similarity		100.0%;	Pred. No. 1;						
Matches 9;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
QY	239 QCAAGCTGP 247								
Db	235 QCAAGCTGP 243								
RESULT 8									
ID	DHCA_RAT	STANDARD;	PRT;	276 AA.					
AC	P47727;								
DT	01-FEB-1996 (Rel. 33, Created)								
DT	01-FEB-1996 (Rel. 33, Last sequence update)								
DT	16-OCT-2001 (Rel. 40, Last annotation update)								
DE	Carboxyl reductase [NADPH] 1 (Ec 1.1.1.184) (NADPH-dependent carboxyl reductase 1).								
GN	CBR1 OR CBR.								
OS	Rattus norvegicus (Rat).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.								
OX	NCBI_Taxid=10116;								
RN	SEQUENCE FROM N.A.								
RP	STRAIN=Sprague-Dawley; TISSUE=Testis;								
RC	MEDLINE=9520378; PubMed=7705364;								
RX	Wermuth B., Maeder-Helmenann G., Ernst E.;								
RT	"Cloning and expression of carboxyl reductase from rat testis."								
RL	Eur. J. Biochem. 228:473-479(1995).								
CC	-1- SUBUNIT ACTIVITY: R-CHOH-R' + NADP(+) = R-CO-R' + NADPH.								
CC	-1- CATALYTIC ACTIVITY: (BY SIMILARITY).								
CC	-1- SUBUNIT: MONOMER (BY SIMILARITY).								
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.								
CC	-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (GSR) FAMILY.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).								
CC	-----								
DR	EMBL; X84349; CAA59088.1; -.								
DR	EMBL; X95986; CAA65230.1; -.								
DR	HSSP; P50162; IAE1.								
DR	InterPro; IPR002198; ADH_short.								
DR	Pfam; PF00106; adh_short.1.								
DR	PROSITE; PS00061; ADH_SHORT; 1.								
KW	Oxidoreductase; NADP; Acetylation.								
FT	INT_MET 0								
FT	MOD_RES 0								
FT	NP_BIND 1								
FT	ACT_SITE 9								
FT	ACT_SITE 193								
SO	SEQUENCE 276 AA;	30447 MW;	99B9E77C5E2922AB CRC64;						
Query Match 1.9%;									
Score 8;									

	Best Local Similarity	100.0%	Pred. No.	2.9	Matches	8:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
OY	13 LALLPPGA 20													
Ddb	253 LALLPPGA 260													
	RESULT: 9													
ID	YV30.MYCITU	STANDARD;	PRT;	463 AA.										
AC	007035;													
DT	30-MAY-2000 (Rel. 39, Created)													
DT	30-MAY-2000 (Rel. 39, Last sequence update)													
DE	15-JUN-2002 (Rel. 41, Last annotation update)													
CD	Hypothetical protein RY3130C													
CN	R3130C OR MT3216 OR MTCY03A2.28 OR MTCY164.41c.													
OS	Mycobacterium tuberculosis.													
OC	Bacteria; Actinobacteria (class); Actinobacteridae;													
OX	Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.													
NCBI_Taxid=1773;	[1]													
RP	SEQUENCE FROM N.A.													
RC	STRAIN-H37RV.													
RA	MEDLINE=98295987; PubMed=9634230;													
RA	Cole S.T., Brosch R., Parkhill T., Garnier T., Churcher C., Harris D.,													
RA	Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,													
RA	Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,													
RA	Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,													
RA	Honrsy T., Tagels K., Krogh A., McLean J., Koule S., Murphy L.,													
RA	Oliver S., Osborne J., Quail M.A., Rajadream M.A., Rogers J.,													
RA	Rutten S., Seeger K., Skelton S., Squares S., Squares R.,													
RA	Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;													
RT	"Deciphering the biology of Mycobacterium tuberculosis from the													
RL	complete genome sequence."													
Nature 393:537-544(1998).	[2]													
RM	SEQUENCE FROM N.A.													
RP	STRAIN-CDC 1551 / Oshkosh;													
RA	Fleischmann R.D., Alland A., Eisen J.A., Carpenter L., White O.,													
RA	Petersen J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,													
RA	Kolony J.F., Nelson W.C., Umayam L.A., Encmlava M.D., Salzberg S.L.,													
RA	Delcher A., Utterback T., Weidman J., Knout H., Gill J., Mikula A.,													
RA	Bisnal W.;													
RT	"Whole genome comparison of Mycobacterium tuberculosis clinical and													
RT	laboratory strains."													
Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.	-1-													
CC	1- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.													
CC	-----													
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration													
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -													
CC	the European Bioinformatics Institute. There are no restrictions on its													
CC	use by non-profit institutions as long as its content is in no way													
CC	modified and this statement is not removed. Usage by and for commercial													
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/													
CC	or send an email to license@lsb.sib.ch).													
CC	-----				</									

Db 315 PLOGLRIV 322

RESULT 10
NCA2_MOUSE
ID NCA2_MOUSE STANDARD: PRT: 725 AA.

AC P13594; 061950; (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120) (NCAM-120).

GN NCAM1 OR NCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=87246524; PubMed=3595563;
RA Barthels D., Santoni M.-J., Wille W., Ruppert C., Caix J.-C.,
RA Hirsch M.-R., Pontecilla-Camps J.-C., Goridis C.;
RT "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for
a Mr 79,000 polypeptide without a membrane-spanning region.";
RL EMO J. 6:907-914(1987).
RN [2]
RP SEQUENCE OF 20-700 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=89251563; PubMed=2721486;
RA Santoni M.-J., Barthels D., Vopper G., Boned A., Goridis C., Wille W.;
RT "Different exon usage involving an unusual splicing mechanism
generates at least eight types of NCAM cDNA in mouse brain.";
RL EMO J. 8:385-392(1989).
RN [3]
RP SEQUENCE OF 642-725 FROM N.A.
RX MEDLINE=88283628; PubMed=3396534;
RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;
RT "Different splicing and alternative polyadenylation generates
distinct NCAM transcripts and proteins in the mouse.";
RL EMO J. 7:625-632(1988).
RN [4]
RP SEQUENCE OF 20-36.
RX MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
domain of mammalian neural cell adhesion molecules.";
RL J. Biol. Chem. 261:3396-3401(1986).
CC -1- FUNCTION: THIS PROTEIN IS A CELL ADHESION MOLECULE INVOLVED IN
CC NEURON-NEURON ADHESION, NEURITE FASCICULATION, OUTGROWTH OF
CC NEURITES, ETC.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms: N-CAM 180 (AC P13595), N-CAM 140
CC (AC P13595) and N-CAM 120 (shown here); are produced by
CC alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Y00051; CAAB6263.1; -;
DR EMBL: X15049; CAA33148.1; ALT_SEQ.
DR EMBL: X07195; CAA30173.1; -;
DR PIR: A29673; IJMSNG.
DR MGD: MGI:97281; Ncam1.
DR InterPro: IPR003961; FN_III.

DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_c2.
DR Pfam: PF00041; fn3; 2.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00060; FN3; 2.
DR SMART: SM00408; IGC2; 5.
KW Cell adhesion; Glycoprotein; Repeat; Alternative splicing;
KW Immunoglobulin domain; Signal; Heparin-binding; GPI-anchor.
FT SIGNAL 1 19
FT CHAIN 20 725
FT
FT DOMAIN 34 103
FT DOMAIN 132 196
FT DOMAIN 228 295
FT DOMAIN 323 393
FT DOMAIN 420 487
FT DOMAIN 519 596
FT DOMAIN 625 692
FT DOMAIN 152 156
FT DOMAIN 161 165
FT DISULFD 41 96
FT DISULFD 139 189
FT DISULFD 235 288
FT DISULFD 330 386
FT DISULFD 427 480
FT CARBOHD 316 316
FT CARBOHD 348 348
FT CARBOHD 424 424
FT CARBOHD 450 450
FT CARBOHD 479 479
FT CARBOHD 261 268
FT CONFLICT 273 273
FT CONFLICT 354 355
FT CONFLICT 549 549
FT CONFLICT 572 572
FT CONFLICT 575 575
FT CONFLICT 589 594
FT CONFLICT 600 602
FT CONFLICT 657 657
SQ SEQUENCE 725 AA; 80296 MW; C2AEBB84461C6B2F CAC64;

Query Match 1.98; Score 8; DB 1; Length 725;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 322 VTAEDGTQ 329
Db 98 VTAEDGTQ 105

RESULT 11
NCA1_RAT
ID NCA1_RAT STANDARD: PRT: 858 AA.

AC P13596;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Neural cell adhesion molecule 1, 140 kDa isoform precursor (N-CAM 140) (NCAM-140).

GN NCAM1 OR NCAM.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=88059265; PubMed=3680385;
RA Small S.J., Shull G.E., Santoni M.-J., Akesson R.;
RT "Identification of a cDNA clone that contains the complete coding
sequence for a 140-kD rat NCAM polypeptide.";
RL J. Cell Biol. 105:2335-2345(1987).

CC -I- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X07200; CAA30177.1; -;
 DR EMBL: Y00051; -; NOT_ANNOTATED_CDS.
 DR EMBL: X06328; CAA29641.1; -;
 DR EMBL: X07195; CAA30173.1; -;
 DR EMBL: X07244; CAA30230.1; -;
 DR EMBL: X15051; CAA33150.1; -;
 DR EMBL: X15052; CAA33151.1; -;
 DR PIR: A29673; IJMSNL.
 DR MGD; MGI:97281; Ncam1.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR Pfam: PF00041; fn3; 2.
 DR Pfam: PF00047; Ig; 5; 2.
 DR SMART; SM0060; FN3; 2.
 DR SMART; SM00408; IGC2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Repeat;
 KW Immunoglobulin domain; Alternative splicing; Signal; Heparin-binding.
 FT SIGNAL 1 19
 FT CHAIN 20 1115
 FT TRANSSEM 20 711
 FT DOMAIN 712 729
 FT DOMAIN 730 1115
 FT DOMAIN 34 103
 FT DOMAIN 132 196
 FT DOMAIN 228 295
 FT DOMAIN 323 393
 FT DOMAIN 420 487
 FT DOMAIN 519 596
 FT DOMAIN 625 692
 FT DOMAIN 152 156
 FT DOMAIN 161 165
 FT DISULFID 41 96
 FT DISULFID 139 189
 FT DISULFID 235 288
 FT DISULFID 330 386
 FT DISULFID 427 480
 FT CARBOHYD 222 222
 FT CARBOHYD 316 316
 FT CARBOHYD 348 348
 FT CARBOHYD 424 424
 FT CARBOHYD 450 450
 FT CARBOHYD 479 479
 FT VARSPLIT 810 1076
 SQ SEQUENCE 1115 AA; 119351 MW; 2C93DCD474CFB9AF CRC64;
 Query Match 1.7%; Score 8; DB 1; Length 1115;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 322 VTAEDEGTQ 329
 Db 98 VTAEDEGTQ 105
 RESULT 13
 SECE_SULSO STANDARD; PRT; 62 AA.
 ID SECE_SULSO
 AC P58191;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Preprotein translocase sece subunit (Protein transport protein SEC61
 DE gamma subunit homolog).
 DE SECE OR SSO5663.
 OS Sulfolobus solfataricus.
 OC Archaeae; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Comtatlonieri F., Zivanovic Y., Allard G.,
 RA Aveyez M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -I- FUNCTION: Involved in protein export (By similarity).
 CC -I- SUBUNIT: Component of the protein translocase complex (By
 CC similarity).
 CC -I- SUBCELLULAR LOCATION: Tail-anchored membrane protein (By
 CC similarity).
 CC -I- SIMILARITY: BELONGS TO THE SECE/SEC61-GAMMA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE006668; AAK0677.1; -;
 DR InterPro: IPR001901; SECE.
 DR InterPro: IPR004795; SECE_euk_arch.
 DR TIGRFAMs: TIGR00327; SECE_euk_arch; 1.
 DR PROSITE: PS01067; SECE_SEC61G; FALSE_NEG.
 KW Protein transport; Translocation; Transmembrane; Complete proteome.
 FT TRANSSEM 40 60
 FT SEQUENCE 62 AA; 7184 MW; CD51874P424B99FA CRC64;
 Query Match 1.7%; Score 7; DB 1; Length 62;
 Best Local Similarity 100.0%; Pred. No. 8.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 178 QLATFLI 184
 Db 54 QLATFLI 60
 RESULT 14
 NADE_HUMAN STANDARD; PRT; 111 AA.
 ID NADE_HUMAN
 AC Q00994;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE p75NTR-associated cell death executor (Nerve growth factor receptor
 DE associated protein 1) (Ovarian granulosa cell 13.0 kDa protein HGR74).
 GN NGFRAP1 OR NADE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RX MEDLINE=91025550; PubMed=2171551;
 RA Rapp G., Freudenstein J., Klaudiny J., Mucha J., Wempe F., Zimmer M.,
 RA Scheit K.H.;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 17:11:37 ; Search time 24.3996 Seconds

(without alignments)
1650.860 Million cell updates/sec

Title: US-09-506-079E-2

Perfect score: 419

Sequence: 1 METALCKRWGLLALLIPGA.....VGRKDPDAHVAVLSRYEG 419

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	83	19.8	1255 1 A24571	protein-tyrosine k
2	51	12.2	1254 1 I48161	p-185 precursor -
3	51	12.2	1260 1 TVRPNV	protein-tyrosine k
4	9	2.1	230 2 B82252	knife-related prote
5	9	2.1	527 2 A42032	epidermal growth f
6	9	2.1	644 2 A36325	epidermal growth f
7	9	2.1	1210 1 GOHUB	epidermal growth f
8	9	2.1	1210 2 A53183	epidermal growth f
9	9	2.1	1223 1 TVCHLV	epidermal growth f
10	8	1.9	277 2 UC5284	carbonyl reductase
11	8	1.9	341 2 B83298	conserved hypotet
12	8	1.9	463 2 H70922	hypothetical prote
13	8	1.9	725 1 IJMSNG	neural cell adhesi
14	8	1.9	858 1 IJRTNG	neural cell adhesi
15	8	1.9	1115 1 IJMSNL	neural cell adhesi
16	7	1.7	62 2 F90177	hypothetical prote
17	7	1.7	111 2 C35826	hypothetical 13k p
18	7	1.7	121 2 AH1848	hypothetical prote
19	7	1.7	171 2 F97564	(Y09560) mobb prot
20	7	1.7	171 2 AE2785	hypothetical prote
21	7	1.7	173 2 SI0199	NMDH2 dehydrogenas
22	7	1.7	209 2 T00733	hypothetical prote
23	7	1.7	220 2 S05595	trypsin inhibitor
24	7	1.7	220 2 I50588	fibroblast growth
25	7	1.7	221 2 JC7587	stromal cell-deriv
26	7	1.7	223 2 H83462	heme exporter prot
27	7	1.7	224 1 A33861	trans-activating t
28	7	1.7	248 1 C83431	type III export pr
29	7	1.7	259 2 D75275	endonuclease III -

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth
A:Reference number: A24571; MUID:8611663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Cousens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALU', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M11730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: GB:I29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Pal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcrip
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:9183983; PIDN:AAA58637.1; PID:9553332
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C:Genetics:
 A:Gene: GDB:ERBB2; NGL; NEU; HER-2
 A:Cross-references: GDB:120613; OMIM:164870
 A:Map position: 17q21.1-17q21.1
 A:Introns: 25/1, 75/3, 147/1, 863/3
 A:Note: the list of introns is incomplete
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-125/Domain: protein-tyrosine kinase erbB2 #status predicted <MNT>
 F:122-653/Domain: extracellular #status predicted <EXT>
 F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
 F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
 F:654-675/Domain: transmembrane #status predicted <TM>
 F:676-1255/Domain: intracellular #status predicted <INT>
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif
 F:686-124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:753/Active site: Lys #status predicted
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 19.8%; Score 83; DB 1; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 2.9e-74;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 CAGSRWGSSEDCOSLRTVCAGCARGKPLPTDCCHBQCAAGTGPKNHSDCLALHF 258
 DB 199 CAGSRWGSSEDCOSLRTVCAGCARGKPLPTDCCHBQCAAGTGPKNHSDCLALHF 258

QY 259 NMSGICELCPALVTYNTDFES 281
 DB 259 NMSGICELCPALVTYNTDFES 281

RESULT 2
 148161
 P:185 precursor - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C:Accession: 148161
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa, Gene 140, 251-255, 1994
 A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
 A:Reference number: 148161; MUID:94193007; PMID:7908275
 A:Accession: 148161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1254 <RES>
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BAA03801.1; PID:9747595
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 12.2%; Score 51; DB 2; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 3e-42;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LPTDCHEQCAAGCTGPKHSDCLALHFNMSGICELCPALVTYNTDFES 281
 DB 231 LPTDCHEQCAAGCTGPKHSDCLALHFNMSGICELCPALVTYNTDFES 281

RESULT 3
 TVRTND

protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 C:Accession: A24562; A61204
 R:Barthmann, C.I.; Hung, M.C.; Weinberg, R.A.
 Nature 319, 226-230, 1986
 A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein
 A:Reference number: A24562; MUID:86118662; PMID:3945311
 A:Accession: A24562
 A:Molecule type: mRNA
 A:Residues: 1-1260 <BAR>
 A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746
 R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohe Carcinogenesis 12, 1975-1978, 1991
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals a 2-thiazolidylformamide or N-methyl-N-nitrosourea.
 A:Reference number: A61204; MUID:92035293; PMID:1682063
 A:Accession: A61204
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 637-663, 'V', 665-702 <MAS>
 A:Note: authors translated the codon GCA for residue 25 as Val
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1260/Domain: protein-tyrosine kinase neu #status predicted <MNT>
 F:658-680/Domain: transmembrane #status predicted <TM>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:71,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 12.2%; Score 51; DB 1; Length 1260;
 Best Local Similarity 100.0%; Pred. No. 3e-42;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 LPTDCHEQCAAGCTGPKHSDCLALHFNMSGICELCPALVTYNTDFES 281
 DB 235 LPTDCHEQCAAGCTGPKHSDCLALHFNMSGICELCPALVTYNTDFES 285

RESULT 4
 B82252
 R:EF-related protein VC1012 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: B82252
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R. Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bess, S.; Qin, H.; Dragol, I.; Sellers, L., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: B82252
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-230 <HEI>
 A:Cross-references: GB:AE004183; GB:AE003852; NID:9655473; PIDN:AAF94173.1; GSPDB:GN
 A:Experimental source: serogroup O1, strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC1012
 A:Map position: 1
 C:Superfamily: conserved hypothetical protein H11688

Query Match 2.1%; Score 9; DB 2; Length 230;
 Best Local Similarity 100.0%; Pred. No. 0.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LVALPPGA 20

Db 184 LLAUPECA 192

RESULT 5

A:Accession: A42032
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence.revision 31-Dec-1993 #text.change 18-Jun-1999

R:Flückinger, T.W.; Mahlow, N.J.; Kung, H.J.

A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, t

A:Reference number: A42032; MUID:92123214; PMID:1732751

A:Accession: A42032

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-527 <FL>

A:Cross-references: GB:M7637; NID:g211737; PIDN:AAA48759.1; PID:g211738

A:Experimental source: liver

A:Note: sequence extracted from NCBI backbone (NCBIN:76892, NCBI:76893)

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: AMP; growth factor receptor

Query Match

Best Local Similarity 2.1%; Score 9; DB 2; Length 527;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 QCAAGCTCP 247

Db 245 QCAAGCTCP 253

RESULT 6

A:Accession: A36325
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Jan-1991 #sequence.revision 25-Jan-1991 #text.change 10-Oct-1997

C:Accession: A36325

R:Patch, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Earp, H.S.

Mol. Cell. Biol. 10, 2973-2982, 1990

A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded b

A:Reference number: A36325; MUID:90258886; PMID:2342466

A:Accession: A36325

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-644 <PE>

A:Cross-references: GB:M37394

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match

Best Local Similarity 2.1%; Score 9; DB 2; Length 644;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 239 QCAAGCTCP 247

Db 235 QCAAGCTCP 243

RESULT 7

A:Accession: A00641
C:Species: Homo sapiens (man)
C:Date: 15-Nov-1984 #sequence.revision 27-Nov-1985 #text.change 11-Jun-1999

C:Accession: A00641; A25772; S30024; A38672; A43615; A23062; A05281; A60143; A33

R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y

rg, P.H.

Nature 309, 418-425, 1984

A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of

A:Reference number: A00641; MUID:84219729; PMID:6328312

A:Accession: A00641

A:Molecule type: mRNA
A:Residues: 1-1210
A:Cross-references: EMBL:X00588; NID:g31113; PIDN:CAA25240.1; PID:g757924

A:Note: the authors translated the codon AAG for residue 540 as Asn

R:Shih, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Mellino, G.T.; Pastan, I.

Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985

A:Title: Characterization and sequence of the promoter region of the human epidermal

A:Reference number: A25772; MUID:85270438; PMID:2991899

A:Accession: A25772

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-29 <ISH>

A:Cross-references: GB:M1234; NID:g181981; PIDN:AAA52370.1; PID:g553272

R:Haley, J.; Whittle, N.; Bennett, P.; Kitchington, D.; Ullrich, A.; Waterfield, M.

Oncogene Res. 1, 375-396, 1987

A:Title: The human EGF receptor gene: structure of the 110 kb locus and identificatio

A:Reference number: S30024; MUID:88213333; PMID:3329716

A:Accession: S30024

A:Molecule type: DNA

A:Residues: 1-29 <HA2>

A:Cross-references: EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID:g31119

R:Haley, J.D.; Waterfield, M.D.

J. Biol. Chem. 266, 1746-1753, 1991

A:Title: Contributory effects of de Novo transcription and premature transcript termi

A:Reference number: A38672; MUID:91107677; PMID:1988448

A:Accession: A38672

A:Molecule type: DNA

A:Residues: 1-29 <HAL>

A:Cross-references: GB:M38425; NID:g181977; PIDN:AAA63171.1; PID:g553271

A:Experimental source: carcinoma cell line A431-7

R:Xu, Y.; Ishii, S.; Clark, A.O.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;

Nature 309, 806-810, 1984

A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN

A:Reference number: A00642; MUID:84245835; PMID:6330563

A:Accession: A00642

A:Molecule type: mRNA

A:Residues: 'RCAMRA', 150-187, 'KSYQAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-

'798-799, 'TD', 802-811, 'R', 813-942 <XY>

A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF re

R:Lin, C.R.; Chen, W.S.; Kruliger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma,

Science 224, 843-848, 1984

A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplificati

A:Reference number: A43615; MUID:84196372; PMID:6326261

A:Accession: A43615

A:Molecule type: mRNA

A:Residues: 713-964 <LIN>

A:Experimental source: epidermoid carcinoma cell line A431

R:Stamen, F.A.; Gope, M.L.; Schultz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.

Biochem. Biophys. Res. Commun. 124, 125-132, 1984

A:Reference number: A23062; MUID:85046483; PMID:6093780

A:Accession: A23062

A:Molecule type: mRNA

A:Residues: 1028-1210 <STM>

R:Weber, W.; Gill, G.N.; Speiss, J.

Science 224, 294-297, 1984

A:Reference number: A05281; MUID:84172183; PMID:6324343

A:Accession: A05281

A:Molecule type: protein

A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>

R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Scaros, J.V.

J. Biol. Chem. 260, 5205-5208, 1985

A:Title: Identification of residues in the nucleotide binding site of the epidermal g

A:Reference number: A60143; MUID:85182650; PMID:2985580

A:Accession: A60143

A:Molecule type: protein

A:Residues: 740-744, 'X', 746-747 <RUS>

R:Mroczkowski, B.; Mosig, G.; Cohen, S.

Nature 309, 270-273, 1984

A:Title: ATP-stimulated interaction between epidermal growth factor receptor and supe

A:Reference number: A38023; MUID:84191554; PMID:6325948

A:Contents: annotation; receptor activity

A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA

R:Chen, W.S.; Lazar, C.S.; Lund, R.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C

Cell 59, 33-43, 1989
A:Title: Functional independence of the epidermal growth factor receptor from a domain
A:Reference number: A3331; MUID:9000323; PMID:2790960
A:Contents: annotation; internalization signal
C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor
C:Genetics:
A:Gene: GDB:EGFR
A:Cross-references: GDB:120610; OMIM:131550
A:Map position: 7p12.3-7p12.1
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-1210/Product: EGF receptor #status predicted <MAT>
F:25-645/Domain: extracellular #status predicted <EXT>
F:75-300/Domain: EGF receptor extracellular domain repeat <EE1>
F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>
F:646-668/Domain: transmembrane #status predicted <TM>
F:669-1210/Domain: intracellular #status predicted <INT>
F:710-975/Domain: protein kinase homology <KIN>
F:718-726/Region: protein kinase ATP-binding motif
F:999-1046/Region: coated-pit mediated internalization signal
F:1047-1210/Region: inhibitory
F:128,175,352,413,444,528,603/Binding site: carbohydrate (asn) (covalent) #status predic
F:745/Active site: Lys #status experimental

Query Match 2.1%; Score 9; DB 1; Length 1210;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 QCAAGCTGP 247
Db 235 QCAAGCTGP 243

RESULT 8
A53183
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.;
Genes Dev. 8, 399-413, 1994
A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
A:Reference number: A53183; MUID:94170986; PMID:812525
A:Accession: A53183
A:Molecule type: mRNA
A:Residues: 1-1210 <LUE>
A:Cross-references: GB:U03425
R:Avitl, A.; Lay, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
Oncogene 6, 673-676, 1991
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding sit
A:Reference number: A43818; MUID:91232866; PMID:2030916
A:Accession: A43818
A:Molecule type: mRNA
A:Residues: 1-714 <AVT>
A:Cross-references: GB:X59698
R:Eisinger, D.P.; Serrero, G.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24942
A:Accession: S24942
A:Molecule type: mRNA
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>
A:Cross-references: EMBL:Z12608
R:Heisermann, G.J.; Gill, G.N.
J. Biol. Chem. 263, 13152-13158, 1988
A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylated
A:Reference number: A28941; MUID:88330814; PMID:3138233
A:Accession: A28941
A:Molecule type: protein
A:Residues: 689-994, 'X', 696-704, 'L', 706-707, 989-992, 'XX', 995-996, 'X', 998-1000, 1002-1009,
R:Hibbs, M.L.; Dunn, A.R.; Alexander, W.S.
submitted to the EMBL Data Library, April 1994
A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor

A:Reference number: S45325
A:Accession: S45325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-971, 'K', 973-1210 <VER>
A:Cross-references: EMBL:X78997; NID:9488830; PIDN:CA55587.1; PID:9488831
R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Day, S.K.
Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
A:Title: Expression of the epidermal growth factor receptor gene is regulated in mous
A:Reference number: I49643; MUID:93126380; PMID:7678348
A:Accession: I49643
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 12-20, 22-132 <RES>
A:Cross-references: GB:I06864; NID:9193001; PIDN:AA53029.1; PID:9567201
C:Genetics:
A:Gene: EGFR
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phospho
F:1-24/Domain: signal sequence #status predicted <SIG>
F:648-670/Domain: transmembrane #status predicted <TM>
F:712-977/Domain: protein kinase homology <KIN>
F:720-728/Region: protein kinase ATP-binding motif
F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental
F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental
F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimenta
F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 2.1%; Score 9; DB 2; Length 1210;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 QCAAGCTGP 247
Db 235 QCAAGCTGP 243

RESULT 9
TYCHIV
C:Species: Gallus gallus (chicken)
N:Contains: Protein tyrosine kinase (EC 2.7.1.112) erbB
C:Date: 28-Feb-1986 #sequence_revision 05-May-1995 #text_change 04-Feb-2000
C:Accession: A27720; A00643
R:Lax, I.; Johnson, A.; Howk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Venn
Mol. Cell. Biol. 8, 1970-1978, 1988
A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in
A:Reference number: A27720; MUID:88261272; PMID:3260329
A:Accession: A27720
A:Molecule type: mRNA
A:Residues: 1-1223 <LAX>
A:Cross-references: GB:M20386
R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines
Cell 41, 719-726, 1985
A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and
A:Reference number: A00643; MUID:85228222; PMID:2988784
A:Accession: A00643
A:Molecule type: mRNA
A:Residues: 585-1223 <NTL>
A:Cross-references: GB:M10066
C:Genetics:
A:Gene: erbB
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-1223/Product: epidermal growth factor receptor #status predicted <MAT>
F:31-654/Domain: extracellular #status predicted <EXT>
F:81-307/Domain: EGF receptor extracellular domain repeat <EE1>
F:397-610/Domain: EGF receptor extracellular domain repeat <EE2>
F:655-677/Domain: transmembrane #status predicted <TM>
F:678-1223/Domain: intracellular #status predicted <INT>

F:719-984/Domain: protein kinase homology <KIN>
 F:727-735/Region: protein kinase ATP-binding motif
 F:136,202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #S
 F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted
 F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:754/Active site: lys #status predicted
 F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match 2.1%; Score 9; DB 1; Length 1223;
 Best Local Similarity 100.0%; Pred. No. 3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 239 QCAAGCTCP 247
 |||||
 Db 242 QCAAGCTCP 250

RESULT 10

JC5284
 carbonyl reductase (NADPH2) (EC 1.1.1.184), inducible - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 16-Apr-1997 #sequence_revision 18-Jul-1997 #text_change 03-Jun-2002
 C/Accession: S68982; PC2234; JC5284; PC2159; S52349
 R/Wermuth, B.; Maeder-Heinmann, G.; Ernst, E.
 Eur. J. Biochem. 228, 473-479, 1995

A/Title: Cloning and expression of carbonyl reductase from rat testis.
 A/Reference number: S68982; MUID:95220378; PMID:7705364
 A/Accession: S68982

A/Status: preliminary
 A/Molecule type: mRNA

A/Residues: 1-277 <NR>

A/Cross-references: EMBL:X84349; NID:9666086; PIDN:CA59088.1; PID:9666087
 R/Toft, E.; Soederstrom, M.; Ahlberg, M.B.; Depierre, J.W.
 Biochem. Biophys. Res. Commun. 201, 149-154, 1994

A/Title: A novel 34kDa glutathione-binding protein in mature rat ovary.
 A/Reference number: PC2159; MUID:94256971; PMID:8198567

A/Accession: PC2234

A/Molecule type: protein

A/Residues: 104-121, 'X', 123-134, 'D', 136-137, 'H' <TOP>

A/Experimental source: ovary

R/Aoki, H.; Okada, T.; Mizutani, T.; Numata, Y.; Minegishi, T.; Miyamoto, K.
 Biochem. Biophys. Res. Commun. 230, 518-523, 1997

A/Title: Identification of two closely related genes, inducible and noninducible carbonyl
 A/Reference number: JC5284; MUID:97167735; PMID:9015353

A/Contents: ovary

A/Accession: JC5284

A/Molecule type: mRNA

A/Residues: 1140, 'GR', 143, 'R', 145-235, 'T', 237-238, 'E', 240-277 <NR>

A/Cross-references: DDBJ:DB9069; NID:91906811; PIDN:BA19007.1; PID:91906812

C/Comment: This enzyme is a cytosolic, monomeric oxidoreductase that catalyzes the NADPH
 C/Genetics:

A/Status: ICR

C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C/Keywords: cytosol; monomer; NADP; oxidoreductase

F:6-185/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 1.9%; Score 8; DB 2; Length 277;
 Best Local Similarity 100.0%; Pred. No. 8.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LALLPQA 20
 |||||
 Db 254 LALLPQA 261

RESULT 11

B83298
 conserved hypothetical protein PA2771 [Imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C/Accession: B83298
 R/Stover, C.K.; Pham, X.Q.; Errvin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Llm

.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: B83298

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-341 <STO>

A/Cross-references: GB:AE004705; GB:AE004091; NID:9948851; PIDN:AG06159.1; GSPDB:GN
 A/Experimental source: strain PA01

C/Genetics:

Query Match 1.9%; Score 8; DB 2; Length 341;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 LQLRSLTE 147
 |||||
 Db 158 LQLRSLTE 165

RESULT 12

H70922
 hypothetical protein RV3130c - Mycobacterium tuberculosis (strain H37RV)
 C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C/Accession: H70922

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998

A/Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A/Reference number: A70500; MUID:98295987; PMID:9654230

A/Accession: H70922

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-463 <COL>

A/Cross-references: GB:295150; GB:AL123456; NID:93250708; PIDN:CAB08399.1; PID:el3010

A/Experimental source: strain H37RV

C/Genetics:

A/Status: RV3130c

Query Match 1.9%; Score 8; DB 2; Length 463;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 PLQRRLRY 102
 |||||
 Db 315 PLQRRLRY 322

RESULT 13

IUMSNG

neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse

N/Alternate names: NCAM-120

C/Species: Mus musculus (house mouse)

C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 31-Dec-2000

C/Accession: A29673; S00382; A44290

R/Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaux, J.C.; Hirsch, M.R.; Fon

EMBO J. 6, 907-914, 1987

A/Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,

A/Reference number: A29673; MUID:87246524; PMID:3355563

A/Accession: A29673

A/Molecule type: mRNA

A/Residues: 1-725 <BAR>

A/Cross-references: EMBL:Y00051; NID:953342; PIDN:CA68263.1; PID:953343

R/Barbs, J.A.; Chaux, J.C.; Steinmetz, M.; Goridis, C.
 EMBO J. 7, 625-632, 1988
 A/Title: Differential splicing and alternative polyadenylation generates distinct NCA
 A/Reference number: S00382; MUID:88283628; PMID:3396534
 A/Accession: S00382

A.Molecule type: DNA
A.Residues: 642-656, 'D', 658-725 <BA2>
A.Cross-references: EMBL:X07195
R.Rougon, G.; Marshak, D.R.
J. Biol. Chem. 261, 3396-3401, 1986
A.Title: Structural and immunological characterization of the amino-terminal domain of NCAM
A.Reference number: M44290; MUID:86140120; PMID:5112556
A.Accession: M44290
A.Molecule type: protein
A.Residues: 20-36 <ROU>
C.Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM molecule.
C.Genetics: Several forms of NCAM are produced by alternative splicing. See also PIR:IGMS
A.Gene: NCAM
A.Map position: 9
A.Introns: 701/1
C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin domain
C:Keywords: alternative splicing; cell adhesion; duplication; heparin binding; membrane
F:1-19/Domain: signal sequence #status predicted <SIG>
F:34-98/Domain: immunoglobulin homology <IMM1>
F:132-191/Domain: immunoglobulin homology <IMM2>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-290/Domain: immunoglobulin homology <IMM3>
F:263-272/Region: NCAM binding #status predicted
F:323-388/Domain: immunoglobulin homology <IMM4>
F:420-482/Domain: immunoglobulin homology <IMM5>
F:519-596/Domain: fibronectin type III repeat homology <FN3A>
F:625-685/Domain: fibronectin type III repeat homology <FN3B>
F:41-96,139-189,235-288,330-396,437-490/Disulfide bonds: #status predicted
F:222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.9%; Score 8; DB 1; Length 725;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 322 VTAEDETQ 329
DB 98 VTAEDETQ 105
|||||

RESULT 14
IJRTNC
neural cell adhesion molecule short domain form precursor - rat
N:Alternate names: NCAM-140
C:Species: Rattus norvegicus (Norway rat)
C.Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 22-Jun-1999
A.Accession: S00846; B37795; I58136
R.Small, S.J.; Shull, G.E.; Santoni, M.J.; Akesson, R.
J. Cell Biol. 105, 2335-2345, 1987
A.Title: Identification of a cDNA clone that contains the complete coding sequence for NCAM
A.Reference number: S00846; MUID:88059265; PMID:3680385
A.Accession: S00846
A.Molecule type: mRNA
A.Residues: 1-858 <SMA>
A.Cross-references: EMBL:X06564
R.Small, S.J.; Akesson, R.
J. Cell Biol. 111, 2089-2096, 1990
A.Title: Expression of the unique NCAM VASE exon is independently regulated in distinct cell types
A.Reference number: A37795; MUID:91035620; PMID:1699951
A.Accession: B37795
A.Molecule type: mRNA
A.Status: preliminary; not compared with conceptual translation
A.Molecule type: mRNA
A.Residues: 340-381 <SM2>
R.Small, S.J.; Haines, S.L.; Akesson, R.A.
Neuron 1, 1007-1017, 1988
A.Title: Polypeptide variation in an NCAM extracellular immunoglobulin-like fold is developmentally regulated
A.Reference number: I58136; MUID:90166485; PMID:2483093
A.Accession: I58136
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 355-364 <RRS>
A.Cross-references: GB:M32611; NID:9205643; PIDN:AAA1679.1; PID:9205644

C.Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM molecule.
C.Genetics: Various forms of NCAM are produced by alternative splicing.
A.Gene: NCAM
C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin domain
C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; membrane
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-858/Product: neural cell adhesion molecule, short domain form #status predicted
F:34-98/Domain: immunoglobulin homology <IMM1>
F:132-191/Domain: immunoglobulin homology <IMM2>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-290/Domain: immunoglobulin homology <IMM3>
F:263-272/Region: NCAM binding #status predicted
F:323-398/Domain: immunoglobulin homology <IMM4>
F:430-492/Domain: immunoglobulin homology <IMM5>
F:529-606/Domain: fibronectin type III repeat homology <FN3A>
F:635-695/Domain: fibronectin type III repeat homology <FN3B>
F:722-739/Domain: transmembrane #status predicted <TM>
F:740-858/Domain: intracellular #status predicted <INT>
F:41-96,139-189,235-288,330-396,437-490/Disulfide bonds: #status predicted
F:222,316,348,434,460,489/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.9%; Score 8; DB 1; Length 858;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 322 VTAEDETQ 329
DB 98 VTAEDETQ 105
|||||

RESULT 15
IIVSNL
neural cell adhesion molecule 1 precursor, long domain splice form - mouse
N:Alternate names: NCAM-180
C:Contains: neural cell adhesion molecule, short domain splice form (NCAM-140)
C:Species: Mus musculus (house mouse)
C.Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text change 31-Dec-2000
A.Accession: A29673; S00844; S00384; A28281; A44290; S00383
R.Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Pon EMBO J. 6, 907-914, 1987
A.Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79, kDa isoform
A.Reference number: A29673; MUID:87246524; PMID:3595563
A.Accession: A29673
A.Molecule type: mRNA
A.Residues: 1-548, 'T', 550-571, 'T', 573-574, 'D', 576-588, 'MQPS', 593, 'S', 595-599, 'P', 601, 'S', 603-604, 'T', 606-607, 'T', 609-610, 'T', 612-613, 'T', 615-616, 'T', 618-619, 'T', 621-622, 'T', 624-625, 'T', 627-628, 'T', 630-631, 'T', 633-634, 'T', 636-637, 'T', 639-640, 'T', 642-643, 'T', 645-646, 'T', 648-649, 'T', 651-652, 'T', 654-655, 'T', 657-658, 'T', 660-661, 'T', 663-664, 'T', 666-667, 'T', 669-670, 'T', 672-673, 'T', 675-676, 'T', 678-679, 'T', 681-682, 'T', 684-685, 'T', 687-688, 'T', 690-691, 'T', 693-694, 'T', 696-697, 'T', 699-700, 'T', 702-703, 'T', 705-706, 'T', 708-709, 'T', 711-712, 'T', 714-715, 'T', 717-718, 'T', 720-721, 'T', 723-724, 'T', 726-727, 'T', 729-730, 'T', 732-733, 'T', 735-736, 'T', 738-739, 'T', 741-742, 'T', 744-745, 'T', 747-748, 'T', 750-751, 'T', 753-754, 'T', 756-757, 'T', 759-760, 'T', 762-763, 'T', 765-766, 'T', 768-769, 'T', 771-772, 'T', 774-775, 'T', 777-778, 'T', 780-781, 'T', 783-784, 'T', 786-787, 'T', 789-790, 'T', 792-793, 'T', 795-796, 'T', 798-799, 'T', 801-802, 'T', 804-805, 'T', 807-808, 'T', 810-811, 'T', 813-814, 'T', 816-817, 'T', 819-820, 'T', 822-823, 'T', 825-826, 'T', 828-829, 'T', 831-832, 'T', 834-835, 'T', 837-838, 'T', 840-841, 'T', 843-844, 'T', 846-847, 'T', 849-850, 'T', 852-853, 'T', 855-856, 'T', 858-859, 'T', 861-862, 'T', 864-865, 'T', 867-868, 'T', 870-871, 'T', 873-874, 'T', 876-877, 'T', 879-880, 'T', 882-883, 'T', 885-886, 'T', 888-889, 'T', 891-892, 'T', 894-895, 'T', 897-898, 'T', 900-901, 'T', 903-904, 'T', 906-907, 'T', 909-910, 'T', 912-913, 'T', 915-916, 'T', 918-919, 'T', 921-922, 'T', 924-925, 'T', 927-928, 'T', 930-931, 'T', 933-934, 'T', 936-937, 'T', 939-940, 'T', 942-943, 'T', 945-946, 'T', 948-949, 'T', 951-952, 'T', 954-955, 'T', 957-958, 'T', 960-961, 'T', 963-964, 'T', 966-967, 'T', 969-970, 'T', 972-973, 'T', 975-976, 'T', 978-979, 'T', 981-982, 'T', 984-985, 'T', 987-988, 'T', 990-991, 'T', 993-994, 'T', 996-997, 'T', 999-1000, 'T', 1002-1003, 'T', 1005-1006, 'T', 1008-1009, 'T', 1011-1012, 'T', 1014-1015, 'T', 1017-1018, 'T', 1020-1021, 'T', 1023-1024, 'T', 1026-1027, 'T', 1029-1030, 'T', 1032-1033, 'T', 1035-1036, 'T', 1038-1039, 'T', 1041-1042, 'T', 1044-1045, 'T', 1047-1048, 'T', 1050-1051, 'T', 1053-1054, 'T', 1056-1057, 'T', 1059-1060, 'T', 1062-1063, 'T', 1065-1066, 'T', 1068-1069, 'T', 1071-1072, 'T', 1074-1075, 'T', 1077-1078, 'T', 1080-1081, 'T', 1083-1084, 'T', 1086-1087, 'T', 1089-1090, 'T', 1092-1093, 'T', 1095-1096, 'T', 1098-1099, 'T', 1101-1102, 'T', 1104-1105, 'T', 1107-1108, 'T', 1110-1111, 'T', 1113-1114, 'T', 1116-1117, 'T', 1119-1120, 'T', 1122-1123, 'T', 1125-1126, 'T', 1128-1129, 'T', 1131-1132, 'T', 1134-1135, 'T', 1137-1138, 'T', 1140-1141, 'T', 1143-1144, 'T', 1146-1147, 'T', 1149-1150, 'T', 1152-1153, 'T', 1155-1156, 'T', 1158-1159, 'T', 1161-1162, 'T', 1164-1165, 'T', 1167-1168, 'T', 1170-1171, 'T', 1173-1174, 'T', 1176-1177, 'T', 1179-1180, 'T', 1182-1183, 'T', 1185-1186, 'T', 1188-1189, 'T', 1191-1192, 'T', 1194-1195, 'T', 1197-1198, 'T', 1200-1201, 'T', 1203-1204, 'T', 1206-1207, 'T', 1209-1210, 'T', 1212-1213, 'T', 1215-1216, 'T', 1218-1219, 'T', 1221-1222, 'T', 1224-1225, 'T', 1227-1228, 'T', 1230-1231, 'T', 1233-1234, 'T', 1236-1237, 'T', 1239-1240, 'T', 1242-1243, 'T', 1245-1246, 'T', 1248-1249, 'T', 1251-1252, 'T', 1254-1255, 'T', 1257-1258, 'T', 1260-1261, 'T', 1263-1264, 'T', 1266-1267, 'T', 1269-1270, 'T', 1272-1273, 'T', 1275-1276, 'T', 1278-1279, 'T', 1281-1282, 'T', 1284-1285, 'T', 1287-1288, 'T', 1290-1291, 'T', 1293-1294, 'T', 1296-1297, 'T', 1299-1300, 'T', 1302-1303, 'T', 1305-1306, 'T', 1308-1309, 'T', 1311-1312, 'T', 1314-1315, 'T', 1317-1318, 'T', 1320-1321, 'T', 1323-1324, 'T', 1326-1327, 'T', 1329-1330, 'T', 1332-1333, 'T', 1335-1336, 'T', 1338-1339, 'T', 1341-1342, 'T', 1344-1345, 'T', 1347-1348, 'T', 1350-1351, 'T', 1353-1354, 'T', 1356-1357, 'T', 1359-1360, 'T', 1362-1363, 'T', 1365-1366, 'T', 1368-1369, 'T', 1371-1372, 'T', 1374-1375, 'T', 1377-1378, 'T', 1380-1381, 'T', 1383-1384, 'T', 1386-1387, 'T', 1389-1390, 'T', 1392-1393, 'T', 1395-1396, 'T', 1398-1399, 'T', 1401-1402, 'T', 1404-1405, 'T', 1407-1408, 'T', 1410-1411, 'T', 1413-1414, 'T', 1416-1417, 'T', 1419-1420, 'T', 1422-1423, 'T', 1425-1426, 'T', 1428-1429, 'T', 1431-1432, 'T', 1434-1435, 'T', 1437-1438, 'T', 1440-1441, 'T', 1443-1444, 'T', 1446-1447, 'T', 1449-1450, 'T', 1452-1453, 'T', 1455-1456, 'T', 1458-1459, 'T', 1461-1462, 'T', 1464-1465, 'T', 1467-1468, 'T', 1470-1471, 'T', 1473-1474, 'T', 1476-1477, 'T', 1479-1480, 'T', 1482-1483, 'T', 1485-1486, 'T', 1488-1489, 'T', 1491-1492, 'T', 1494-1495, 'T', 1497-1498, 'T', 1500-1501, 'T', 1503-1504, 'T', 1506-1507, 'T', 1509-1510, 'T', 1512-1513, 'T', 1515-1516, 'T', 1518-1519, 'T', 1521-1522, 'T', 1524-1525, 'T', 1527-1528, 'T', 1530-1531, 'T', 1533-1534, 'T', 1536-1537, 'T', 1539-1540, 'T', 1542-1543, 'T', 1545-1546, 'T', 1548-1549, 'T', 1551-1552, 'T', 1554-1555, 'T', 1557-1558, 'T', 1560-1561, 'T', 1563-1564, 'T', 1566-1567, 'T', 1569-1570, 'T', 1572-1573, 'T', 1575-1576, 'T', 1578-1579, 'T', 1581-1582, 'T', 1584-1585, 'T', 1587-1588, 'T', 1590-1591, 'T', 1593-1594, 'T', 1596-1597, 'T', 1599-1600, 'T', 1602-1603, 'T', 1605-1606, 'T', 1608-1609, 'T', 1611-1612, 'T', 1614-1615, 'T', 1617-1618, 'T', 1620-1621, 'T', 1623-1624, 'T', 1626-1627, 'T', 1629-1630, 'T', 1632-1633, 'T', 1635-1636, 'T', 1638-1639, 'T', 1641-1642, 'T', 1644-1645, 'T', 1647-1648, 'T', 1650-1651, 'T', 1653-1654, 'T', 1656-1657, 'T', 1659-1660, 'T', 1662-1663, 'T', 1665-1666, 'T', 1668-1669, 'T', 1671-1672, 'T', 1674-1675, 'T', 1677-1678, 'T', 1680-1681, 'T', 1683-1684, 'T', 1686-1687, 'T', 1689-1690, 'T', 1692-1693, 'T', 1695-1696, 'T', 1698-1699, 'T', 1701-1702, 'T', 1704-1705, 'T', 1707-1708, 'T', 1710-1711, 'T', 1713-1714, 'T', 1716-1717, 'T', 1719-1720, 'T', 1722-1723, 'T', 1725-1726, 'T', 1728-1729, 'T', 1731-1732, 'T', 1734-1735, 'T', 1737-1738, 'T', 1740-1741, 'T', 1743-1744, 'T', 1746-1747, 'T', 1749-1750, 'T', 1752-1753, 'T', 1755-1756, 'T', 1758-1759, 'T', 1761-1762, 'T', 1764-1765, 'T', 1767-1768, 'T', 1770-1771, 'T', 1773-1774, 'T', 1776-1777, 'T', 1779-1780, 'T', 1782-1783, 'T', 1785-1786, 'T', 1788-1789, 'T', 1791-1792, 'T', 1794-1795, 'T', 1797-1798, 'T', 1800-1801, 'T', 1803-1804, 'T', 1806-1807, 'T', 1809-1810, 'T', 1812-1813, 'T', 1815-1816, 'T', 1818-1819, 'T', 1821-1822, 'T', 1824-1825, 'T', 1827-1828, 'T', 1830-1831, 'T', 1833-1834, 'T', 1836-1837, 'T', 1839-1840, 'T', 1842-1843, 'T', 1845-1846, 'T', 1848-1849, 'T', 1851-1852, 'T', 1854-1855, 'T', 1857-1858, 'T', 1860-1861, 'T', 1863-1864, 'T', 1866-1867, 'T', 1869-1870, 'T', 1872-1873, 'T', 1875-1876, 'T', 1878-1879, 'T', 1881-1882, 'T', 1884-1885, 'T', 1887-1888, 'T', 1890-1891, 'T', 1893-1894, 'T', 1896-1897, 'T', 1899-1900, 'T', 1902-1903, 'T', 1905-1906, 'T', 1908-1909, 'T', 1911-1912, 'T', 1914-1915, 'T', 1917-1918, 'T', 1920-1921, 'T', 1923-1924, 'T', 1926-1927, 'T', 1929-1930, 'T', 1932-1933, 'T', 1935-1936, 'T', 1938-1939, 'T', 1941-1942, 'T', 1944-1945, 'T', 1947-1948, 'T', 1950-1951, 'T', 1953-1954, 'T', 1956-1957, 'T', 1959-1960, 'T', 1962-1963, 'T', 1965-1966, 'T', 1968-1969, 'T', 1971-1972, 'T', 1974-1975, 'T', 1977-1978, 'T', 1980-1981, 'T', 1983-1984, 'T', 1986-1987, 'T', 1989-1990, 'T', 1992-1993, 'T', 1995-1996, 'T', 1998-1999, 'T', 2001-2002, 'T', 2004-2005, 'T', 2007-2008, 'T', 2010-2011, 'T', 2013-2014, 'T', 2016-2017, 'T', 2019-2020, 'T', 2022-2023, 'T', 2025-2026, 'T', 2028-2029, 'T', 2031-2032, 'T', 2034-2035, 'T', 2037-2038, 'T', 2040-2041, 'T', 2043-2044, 'T', 2046-2047, 'T', 2049-2050, 'T', 2052-2053, 'T', 2055-2056, 'T', 2058-2059, 'T', 2061-2062, 'T', 2064-2065, 'T', 2067-2068, 'T', 2070-2071, 'T', 2073-2074, 'T', 2076-2077, 'T', 2079-2080, 'T', 2082-2083, 'T', 2085-2086, 'T', 2088-2089, 'T', 2091-2092, 'T', 2094-2095, 'T', 2097-2098, 'T', 2100-2101, 'T', 2103-2104, 'T', 2106-2107, 'T', 2109-2110, 'T', 2112-2113, 'T', 2115-2116, 'T', 2118-2119, 'T', 2121-2122, 'T', 2124-2125, 'T', 2127-2128, 'T', 2130-2131, 'T', 2133-2134, 'T', 2136-2137, 'T', 2139-2140, 'T', 2142-2143, 'T', 2145-2146, 'T', 2148-2149, 'T', 2151-2152, 'T', 2154-2155, 'T', 2157-2158, 'T', 2160-2161, 'T', 2163-2164, 'T', 2166-2167, 'T', 2169-2170, 'T', 2172-2173, 'T', 2175-2176, 'T', 2178-2179, 'T', 2181-2182, 'T', 2184-2185, 'T', 2187-2188, 'T', 2190-2191, 'T', 2193-2194, 'T', 2196-2197, 'T', 2199-2200, 'T', 2202-2203, 'T', 2205-2206, 'T', 2208-2209, 'T', 2211-2212, 'T', 2214-2215, 'T', 2217-2218, 'T', 2220-2221, 'T', 2223-2224, 'T', 2226-2227, 'T', 2229-2230, 'T', 2232-2233, 'T', 2235-2236, 'T', 2238-2239, 'T', 2241-2242, 'T', 2244-2245, 'T', 2247-2248, 'T', 2250-2251, 'T', 2253-2254, 'T', 2256-2257, 'T', 2259-2260, 'T', 2262-2263, 'T', 2265-2266, 'T', 2268-2269, 'T', 2271-2272, 'T', 2274-2275, 'T', 2277-2278, 'T', 2280-2281, 'T', 2283-2284, 'T', 2286-2287, 'T', 2289-2290, 'T', 2292-2293, 'T', 2295-2296, 'T', 2298-2299, 'T', 2301-2302, 'T', 2304-2305, 'T', 2307-2308, 'T', 2310-2311, 'T', 2313-2314, 'T', 2316-2317, 'T', 2319-2320, 'T', 2322-2323, 'T', 2325-2326, 'T', 2328-2329, 'T', 2331-2332, 'T', 2334-2335, 'T', 2337-2338, 'T', 2340-2341, 'T', 2343-2344, 'T', 2346-2347, 'T', 2349-2350, 'T', 2352-2353, 'T', 2355-2356, 'T', 2358-2359, 'T', 2361-2362, 'T', 2364-2365, 'T', 2367-2368, 'T', 2370-2371, 'T', 2373-2374, 'T', 2376-2377, 'T', 2379-2380, 'T', 2382-2383, 'T', 2385-2386, 'T', 2388-2389, 'T', 2391-2392, 'T', 2394-2395, 'T', 2397-2398, 'T', 2400-2401, 'T', 2403-2404, 'T', 2406-2407, 'T', 2409-2410, 'T', 2412-2413, 'T', 2415-2416, 'T', 2418-2419, 'T', 2421-2422, 'T', 2424-2425, 'T', 2427-2428, 'T', 2430-2431, 'T', 2433-2434, 'T', 2436-2437, 'T', 2439-2440, 'T', 2442-2443, 'T', 2445-2446, 'T', 2448-2449, 'T', 2451-2452, 'T', 2454-2455, 'T', 2457-2458, 'T', 2460-2461, 'T', 2463-2464, 'T', 2466-2467, 'T', 2469-2470, 'T', 2472-2473, 'T', 2475-2476, 'T', 2478-2479, 'T', 2481-2482, 'T', 2484-2485, 'T', 2487-2488, 'T', 2490-2491, 'T', 2493-2494, 'T', 2496-2497, 'T', 2499-2500, 'T', 2502-2503, 'T', 2505-2506, 'T', 2508-2509, 'T', 2511-2512, 'T', 2514-2515, 'T', 2517-2518, 'T', 2520-2521, 'T', 2523-2524, 'T', 2526-2527, 'T', 2529-2530, 'T', 2532-2533, 'T', 2535-2536, 'T', 2538-2539, 'T', 2541-2542, 'T', 2544-2545, 'T', 2547-2548, 'T', 2550-2551, 'T', 2553-2554, 'T', 2556-2557, 'T', 2559-2560, 'T', 2562-2563, 'T', 2565-2566, 'T', 2568-2569, 'T', 2571-2572, 'T', 2574-2575, 'T', 2577-2578, 'T', 2580-2581, 'T', 2583-2584, 'T', 2586-2587, 'T', 2589-2590, 'T', 2592-2593, 'T', 2595-2596, 'T', 2598-2599, 'T', 2601-2602, 'T', 2604-2605, 'T', 2607-2608, 'T', 2610-2611, 'T', 2613-2614, 'T', 2616-2617, 'T', 2619-2620, 'T', 2622-2623, 'T', 2625-2626, 'T', 2628-2629, 'T', 2631-2632, 'T', 2634-2635, 'T', 2637-2638, 'T', 2640-2641, 'T', 2643-2644, 'T', 2646-2647, 'T', 2649-2650, 'T', 2652-2653, 'T', 2655-2656, 'T', 2658-2659, 'T', 2661-2662, 'T', 2664-2665, 'T', 2667-2668, 'T', 2670-2671, 'T', 2673-2674, 'T', 2676-2677, 'T', 2679-2680, 'T', 2682-2683, 'T', 2685-2686, 'T', 2688-2689, 'T', 2691-2692, 'T', 2694-2695, 'T', 2697-2698, 'T', 2700-2701, 'T', 2703-2704, 'T', 2706-2707, 'T', 2709-2710, 'T', 2712-2713, 'T', 2715-2716, 'T', 2718-2719, 'T', 2721-2722, 'T', 2724-2725, 'T', 2727-2728, 'T', 2730-2731, 'T', 2733-2734, 'T', 2736-27

R; Rougon, G.; Marshak, D.R.
 J. Biol. Chem. 261, 3396-3401, 1986
 A; Title: Structural and immunological characterization of the amino-terminal domain of NCAM
 A; Reference number: A44290; MUID:86140120; PMID:3512556
 A; Accession: A44290
 A; Molecule type: protein
 A; Residues: 20-36 <R0U>
 C; Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM molecule
 C; Comment: Several forms of NCAM are produced by alternative splicing. See also PIR: IJMS
 C; Genetics:
 A; Gene: NCAM
 A; Map position: 9
 A; Introns: 643/3; 701/1; 770/2; 809/2; 1076/2
 C; Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immunoglobulin
 C; Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; signal
 F; 1-19/Domain: signal sequence #status predicted <SIG>
 F; 20-115/Product: neural cell adhesion molecule, long domain splice form #status expert
 F; 20-809,1077-1115/Product: neural cell adhesion molecule, short domain splice form #status expert
 F; 20-711/Domain: extracellular #status predicted <EXT>
 F; 34-98/Domain: immunoglobulin homology <IMM1>
 F; 132-191/Domain: immunoglobulin homology <IMM2>
 F; 152-156/Region: heparin binding #status predicted
 F; 161-165/Region: heparin binding #status predicted
 F; 228-290/Domain: immunoglobulin homology <IMM3>
 F; 262-272/Region: NCAM binding #status predicted
 F; 323-388/Domain: immunoglobulin homology <IMM4>
 F; 420-482/Domain: immunoglobulin homology <IMM5>
 F; 519-596/Domain: fibronectin type III repeat homology <FN3A>
 F; 625-685/Domain: fibronectin type III repeat homology <FN3B>
 F; 712-729/Domain: transmembrane #status predicted <TM>
 F; 730-1115/Domain: intracellular #status predicted <INT>
 F; 41-96,139-189,235-288,330-386,427-480/Dissulfide bonds: #status predicted
 F; 222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.9%; Score 8; DB 1; Length 115;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 VTAEDETQ 329
 |||||
 Db 98 VTAEDETQ 105

Search completed: January 14, 2003, 17:15:20
 Job time : 45.3996 secs

THIS PAGE BLANK (USPTO)